

Best Local Similarity 100.0%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
|||||
DB 48 LESYT 52

RESULT 4
E90199
conserved hypothetical protein (imported) - *Sulfolobus solfataricus*

C:Species: *Sulfolobus solfataricus*

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001

C:Accession: E90199

C:Author: Singh, R.K.; Conlunheth, F.; Zivanovic, Y.; Allard, G.; Aways, M.; Chan, J.; Jettles, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.

C:Title: R.A.; Ragan, M.A.; Sersen, C.W.; Van der Voort, J.

C:Submitted to Genbank: April 2001

C:Description: *Sulfolobus solfataricus* complete genome.

A:Reference number: A99139

A:Accession: E90199

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1133 <RUP>

A:Cross-references: GB:AE00641; MID:312913694; PIRN:AAK40652; 1; GSPDB:GN00155

C:Gene: S800532

C:Superfamily: conserved hypothetical protein MJ1552

Query Match 100.0%; Score 25; DB 2; Length 133;

Best Local Similarity 100.0%; Pred. No. 51;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
|||||

DB 56 LESYT 60

RESULT 4
T02619
hypothetical protein At2g26040 (imported) - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Feb-2001

C:Accession: T02619; 084055

C:Author: R. Kaul, S. Li, X. Kitchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, R.

C:Submitted to the EMBL Data Library: August 1998

C:Description: *Arabidopsis thaliana* chromosome II BAC T19L8 genomic sequence.

A:Reference number: 214681

A:Accession: T02619

A:Status: translated from GR/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1140 <RUP>

A:Cross-references: EMBL:AC004747; NID:9413596; PIR:Q413709

A:Experimental source: cultured Columbia

A:Author: Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.L.; Town, C.D.; Fujii, C.Y.;

M. Koo, H.; Motil, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.

Nature 402: 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: A84420; MID:20083487; PMID:10617197

A:Accession: G84655

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1190 <STOS>

A:Cross-references: GR:AE002093; NID:93413709; FIDN:AC31232.1; GSPDB:GN00139

C:Gene: T19L8.15; AT2g26040

A:Map position: 2

C:Superfamily: *Arabidopsis thaliana* hypothetical protein T1358.30

Query Match 100.0%; Score 25; DB 2; Length 190;

Best Local Similarity 100.0%; Pred. No. 74;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
|||||
DB 146 LESYT 150

RESULT 5
E90187
HAMI protein (imported) - *Sulfolobus solfataricus*

C:Species: *Sulfolobus solfataricus*

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001

C:Accession: E90187

C:Author: Singh, R.K.; Conlunheth, F.; Zivanovic, Y.; Allard, G.; Aways, M.; Chan, J.; Jettles, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.

C:Title: R.A.; Ragan, M.A.; Sersen, C.W.; Van der Voort, J.

C:Submitted to Genbank: April 2001

C:Description: *Sulfolobus solfataricus* complete genome.

A:Reference number: A99139

A:Accession: E90187

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1142 <RUP>

A:Cross-references: GB:AE00641; MID:312913694; PIRN:AAK40756; 1; GSPDB:GN00155

C:Gene: S800432

C:Superfamily: methanococcus jannaschii conserved hypothetical protein MJ0226

Query Match 100.0%; Score 25; DB 2; Length 192;

Best Local Similarity 100.0%; Pred. No. 75;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
|||||

DB 185 LESYT 189

RESULT 6
AH2547
hypothetical protein alr7667 (imported) - *Nostoc* sp. (strain PCC 7120) plasmid pCC7

A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C:Species: *Nostoc* sp.

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C:Accession: AH2547

C:Author: R. Kaneko, N.; Nakamura, Y.; Waki, C.P.; Kurita, T.; Sasamoto, S.; Watanabe, A.; Iri

Nakazaki, T.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Taka

DNA Res. 8: 205-213, 2001

A:Title: Complete genome sequence of the filamentous nitrogen-fixing cyanobacteria

A:Reference number: AB1807; MID:21595285; PMID:11759840

A:Accession: AH2547

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1215 <RUP>

A:Cross-references: GR:AP003607; PIRN:BAW7410.1; PIR:Q1714752; GSPDB:GN00181

A:Experimental source: strain PCC 7120

C:Gene: alr7667

A:Accession: alr7667

A:Map position: 2

C:Superfamily: *alr7667*

Query Match 100.0%; Score 25; DB 2; Length 216;

Best Local Similarity 100.0%; Pred. No. 85;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
|||||

DB 147 LESYT 151

RESULT 7
S14281
ES43 protein - barley

C:Species: *Hordeum vulgare* (barley)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Mar-1999

GenCore version 5.1.4-PS-4578
Copyright (c) 1993 2003 CompuGen Ltd.

3M Protein - protein search using sw model

Run on: April 8, 2003, 11:59:15 Search time 22.943 seconds
(without alignments)
21.766 Million cell updates/sec

Title: US-09-647-749A-1

Perfect score: 25

Sequence: 1 LEVY 5

Scoring table: RLSUM2
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 9613422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: PIR_733*
2: PIR1*
3: PIR2*
4: PIR3*
5: PIR4*

Prod No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	99	2 JG5295	monocyte chemotact
2	25	100.0	99	2 JG2417	monocyte chemotact
3	25	100.0	133	2 E90199	conserved hypotet
4	25	100.0	140	2 T02619	hypothetical prote
5	25	100.0	192	2 E90187	HAM1 protein (limp
6	25	100.0	216	2 AH2547	hypothetical prote
7	25	100.0	227	2 S44281	ES43 protein - bar
8	25	100.0	245	2 G72484	hypothetical prote
9	25	100.0	247	2 B71827	probable 1-acylaly
10	25	100.0	240	2 D64688	probable 1-acylaly
11	25	100.0	302	2 G86920	probable integral
12	25	100.0	382	2 B70888	hypothetical prote
13	25	100.0	372	2 G64388	hypothetical prote
14	25	100.0	354	2 A96596	hypothetical prote
15	25	100.0	379	2 T08277	catenoid biosynt
16	25	100.0	366	2 F87647	phosphoserine amin
17	25	100.0	454	2 T39297	probable atpase -
18	25	100.0	458	2 E97349	membrane associate
19	25	100.0	442	2 G01804	interleukin 3-reou
20	25	100.0	470	2 D41977	retinoid receptor
21	25	100.0	532	2 D71267	probable UDP-N-ace
22	25	100.0	548	2 T16082	hypothetical prote
23	25	100.0	597	2 D62761	lipase/esterase MF
24	25	100.0	634	2 A25646	dnak-type molecula
25	25	100.0	633	2 A55715	dnak-type molecula
26	25	100.0	739	2 B86606	primosomal protein
27	25	100.0	749	2 C72018	primosomal protein
28	25	100.0	753	2 R71472	probable primosoma
29	25	100.0	753	2 D81734	primosomal protein

30	25	100.0	756	2 S47556	LMC II protein -
31	25	100.0	872	2 B57380	RNA-directed RNA p
32	25	100.0	922	2 A01827	matroilipolytreha
33	25	100.0	1157	2 B69163	DNA helicase relat
34	25	100.0	1252	2 S77037	hypothetical prote
35	25	100.0	1679	2 T30271	surfactant protein
36	25	100.0	2004	2 A03314	probable membrane
37	25	100.0	2397	2 A55335	protein precursor
38	25	100.0	2409	2 A66979	versican precursor
39	25	100.0	2543	2 F69679	polyketide synthas
40	25	100.0	2871	2 A55624	thirillin-1 precur
41	25	100.0	5825	2 T1217	polyprotein - fava
42	23	92.0	51	2 B70141	hypothetical prote
43	23	92.0	112	2 A80661	hypothetical prote
44	23	92.0	117	2 T08158	hypothetical lumin
45	23	92.0	119	2 T08271	probable thiorodox

ALIGNMENTS

RESULT 1

JG5295

monocyte chemotactic protein-2 precursor - human

C:Species: Homo sapiens (man)

C:Date: 02-May-1997 #sequence: JG5295 19-791-1997 #rev: _change 20-Jan-2000

C:Accession: JG5295

R:Van Collie, E.; Froyen, G.; Namiyama, H.; Miura, K.; Piten, P.; Van Aelst, L.; V

Blochem. Biophys. Res. Commun. 211, 720-730, 1997

A:Title: Human monocyte chemotactic protein-2: cDNA cloning and regulated expression

A:Reference number: JG5295; M0109224420; PMID:9070881

A:Accession: JG5295

A:Molecule type: mRNA

A:Residues: 1-99 (VAV)

A:Cross-references: GB:Y10802; NID:01924937; PDB:CAA71760.1; PDB:01924938

A:Experimental source: bone marrow

A:Comments: This protein belongs to the beta-chemokine family which is one of the ma

tis and in tumor biology, and contribute to the trafficking and recruitment of the

C:Genetics:

A:Gene: mcp-2

C:Superfamily: macrophage inflammatory protein

F1-23/Domain: signal sequence #status predicted <SIG>

F1-24-96/Product: monocyte chemotactic protein-2 #status predicted <MAT>

Query Match 100.0% Score 25; DB 2; Length 99;

Best local similarity 100.0% Prod No. 37;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 48 LEVY 52

1 LEVY 5

|||||

Db 48 LEVY 52

RESULT 2

JG2417

monocyte chemotactic protein-2 precursor - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 24-Feb-1995 #sequence: JG2417 24-Feb-1995 #text_change: 16-Jul-1999

C:Accession: JG2417

R:Hosang, K.; Knöke, I.; Klading, J.; Meme, E.; Wilke, W.; Schmitt, K.H.

Biochem. Biophys. Res. Commun. 206, 148-151, 1994

A:Title: Forcine leucal cells express monocyte chemotactic protein-2 (MCP-2): At

A:Reference number: JG2417; M0109504716; PMID:7594015

A:Accession: JG2417

A:Molecule type: mRNA

A:Residues: 1-99 (HVS)

A:Cross-references: GB:C49480; NID:01924938; PDB:CAA8471.1; PDB:0684719

A:Experimental source: corpus luteum

A:Superfamily: macrophage inflammatory protein

F1-23/Domain: signal sequence #status predicted <SIG>

F1-24-59/Product: monocyte chemotactic protein-2 #status predicted <MAT>

Query Match 100.0% Score 25; DB 2; Length 99;

PT New peptides useful for inhibiting human immunodeficiency virus type 1
 PT (HIV-1) gp120 induced neuronal cell death
 XX
 XX Claim 1: Page 11; 16pp; English.
 XX
 CC AAY49695 and AAY49696 represent peptides which inhibit HIV-1 gp120
 CC induced neuronal cell death. Pharmaceutical compositions containing the
 CC peptides are useful for treating symptoms caused by neuronal cell loss.
 CC Such conditions especially associated with HIV infection include
 CC encephalopathies, neuropathies, memory loss, dementia, depression,
 CC psychosis and opportunistic infections. The peptides act as antagonists
 CC of HIV-mediated neurotoxicity and subsequent neuronal degeneration.
 CC This enables therapeutic treatment of HIV infection and other
 CC inflammatory neurological diseases, including multiple sclerosis,
 CC tropical spastic paraparesis and Alzheimer's disease.
 XX
 SQ Sequence 5 AA:
 Query Match 100.0%; Score 25; DB 20; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LESTY 5
 DB 1 LESTY 5
 II III
 RESULT 2
 ID AAY56783
 ID AAY56783 standard; Protein; 50 AA.
 XX
 AC AAY56783:
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #17679.
 XX
 XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KM dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN W0200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 20-APR-2001; 2001WC-US12865.
 XX
 21-APR-2001; 2000US-19947P.
 PR 07-JUN-2000; 2000US-02841P.
 PR 07-JUN-2000; 2000US-215747P.
 XX
 PA (CORI-) CORIAX CORP.
 XX
 PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'Alaisonneuve J, Zhang Y, Jen S, Carter D;
 PI WPI; 2001-616774/71.
 DR N-PSDB: AAS59578.
 XX
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 XX Example 1: SEQ ID NO 17978; 1059pp; English.
 PS
 CC Sequences AAY39105-AAY68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis

CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence. For example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the plotted
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp://wipo.int/pub/published_pat_sequences.
 XX
 SQ Sequence 50 AA:
 Query Match 100.0%; Score 25; DB 22; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LESTY 5
 DB 14 LESTY 18
 II III
 RESULT 3
 ID AAM64554
 ID AAM64554 standard; Protein; 68 AA.
 XX
 AC AAM64554:
 XX
 DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe encoded protein SE2 ID NO: 36659.
 XX
 KW Human; brain expressed exon; gene expression analysis; probe;
 KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KM epilepsy; cancer.
 XX
 OS Homo sapiens.
 XX
 PN W0200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 30-JAN-2001; 2001WO-US06647.
 XX
 04-FEB-2001; 2000US-018612.
 PR 20-MAY-2000; 2000US-020156.
 PR 08-JUN-2000; 2000US-060818.
 PR 04-AUG-2000; 2000US-063446.
 PR 21-SEP-2000; 2000US-023407.
 PR 27-SEP-2000; 2000US-023600.
 PR 04-OCT-2000; 2000US-002464.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-483446/52.
 DR
 XX
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX
 XX Example 4: SEQ ID NO: 36659; 650pp + Sequence Listing; English.
 PS
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.

XX
 SQ Sequence 68 AA:

Query Match 100.0% Score 25; DB 22; Length 68;
 Best Local Similarity 100.0% Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESTY 5
 DB 25 LESTY 29

RESULT 4
 AAY07234
 ID AAY07234 standard; Protein: 71 AA.

XX AAY07234:

XX 06-JUL-1999 (first entry)

XX Truncated monocyte chemotactic protein 2 (6-76)

XX KM Wild type: C-C chemokine; monocyte chemotactic protein 2; MCP2; HIV;
 XX KM regulated on activation normal T-cell expressed and secreted; RANTES;
 XX KM truncation; antagonist; medicaments; diagnosis; inflammation; infection;
 XX KM tumour; angiogenesis; hematopoiesis; autoimmune disease; atherosclerosis;
 XX KM pulmonary disease; skin disorder.

XX OS Homo sapiens.
 XX OS Synthetic.

XX PN EP906524-A1.

XX PD 07-APR-1999.

XX PF 29-SEP-1997; 97EP-0116863.

XX PR 29-SEP-1997; 97EP-0116863.

XX PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV

XX PC Proost P, Struyt S, Van Damme J.

XX ER WPI: 1999-20709/18.

XX PT New amino-terminally truncated C-C chemokines have antagonistic
 XX PT activity for treatment of immune, inflammatory and infectious
 XX PT diseases

XX PS Claim 4; Fig 1; 29pp; English.

XX This sequence represents a truncated C-C chemokine monocyte chemotactic
 XX protein 2 (MCP2) containing amino acids 6-76 of the mature protein.
 XX The invention relates the generation of amino-terminal truncated C-C
 XX chemokines, having chemokine antagonistic activity. The new chemokines
 XX are useful as medicaments, for diagnosis and/or treatment of diseases
 XX which require antagonistic activity of a chemokine e.g. inflammatory
 XX diseases, HIV infection, tumours, and angiogenesis and hematopoiesis
 XX related diseases, including auto-immune diseases, atherosclerosis,
 XX pulmonary diseases and skin disorders.

XX SQ Sequence 71 AA:

Query Match 100.0% Score 25; DB 20; Length 71;

Best Local Similarity 100.0% Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESTY 5
 DB 29 LESTY 24

RESULT 5
 AAY07238
 ID AAY07238 standard; Protein: 71 AA.

XX AAY07238:

XX 06-JUL-1999 (first entry)

XX Truncated monocyte chemotactic protein 2 (6-76).

XX KM Wild type: C-C chemokine; monocyte chemotactic protein 2; MCP2; HIV;
 XX KM regulated on activation normal T-cell expressed and secreted; RANTES;
 XX KM truncation; antagonist; medicaments; diagnosis; inflammation; infection;
 XX KM tumour; angiogenesis; hematopoiesis; autoimmune disease; atherosclerosis;
 XX KM pulmonary disease; skin disorder.

XX OS Homo sapiens.
 XX OS Synthetic.

XX PN EP9065241-A1.

XX PD 31-MAR-1999.

XX PE 10-MAR-1998; 98EP-0104216.

XX PF 19-DEC-1997; 97EP-0122471.

XX PR 29-SEP-1997; 97EP-0116863.

XX PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.

XX PC Proost P, Struyt S, Van Damme J.

XX ER WPI: 1999-206774/18.

XX PT New amino-terminally truncated C-C chemokines have antagonistic
 XX PT activity for treatment of immune, inflammatory and infectious
 XX PT diseases

XX PS Claim 4; Fig 1; 29pp; English.

XX This sequence represents a truncated C-C chemokine monocyte chemotactic
 XX protein 2 (MCP2) containing amino acids 6-76 of the mature protein.
 XX The invention relates the generation of amino-terminal truncated C-C
 XX chemokines, having chemokine antagonistic activity. The new chemokines
 XX are useful as medicaments, for diagnosis and/or treatment of diseases
 XX which require antagonistic activity of a chemokine e.g. inflammatory
 XX diseases, HIV infection, tumours, and angiogenesis and hematopoiesis-
 XX related diseases, including auto-immune diseases, atherosclerosis,
 XX pulmonary diseases and skin disorders.

XX SQ Sequence 71 AA:

Query Match 100.0% Score 25; DB 20; Length 71;

Best Local Similarity 100.0% Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESTY 5
 DB 29 LESTY 24

RESULT 6
 AAR70804
 ID AAR70804 standard; Protein: 72 AA.

XX AAR70804:

XX 29-AUG-1995 (first entry)

XX Chemottractant MCP-2.

XX Chemottractant; MCP-2; heparanase; heparin; heparan sulfate;

KW arthritis; restenosis; cancer; wound healing.
 XX Homo sapiens.
 OS W09504158-A.
 PN 09-FEB-1995.
 PD 26-JUL-1994: 94MO-US08207.
 PF 29-JUL-1993: 94US-0099866.
 PR 13-OCT-1993: 93US-0136117.
 XX (UPJO) UPJOHN CO.
 PA Hooywerf AJ, Ledbetter SR:
 PI WPI: 1995-092239/11.
 DR Screening for c-fos with anti-heparanase activity - by detecting
 XX inhibition of heparin or heparan sulphate degradation, cancer.
 PT potentially useful for treating arthritis, restenosis, cancer.
 PS Claim 13: Page 53: 60pp; English.
 CC Puriified heparanases, prepared under reducing conditions and
 CC activated with transglutaminase, are given in AAR70786-804. Most
 CC are prepared by reverse transcription of mRNA from activated human
 CC leukocytes, then cloning of the cDNA into pVL1392 baculovirus
 CC vector and expression in Sf9 cells in the presence of reduced
 CC glutathione and dithiothreitol.
 XX
 SQ Sequence 72 AA:
 Query Match 100.0%; Score 25; DB 16; Length 72;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LEYSY 5
 DB 24 LEYSY 27
 RESULT 7
 AAAY9031
 ID AAAY9031 standard; Protein: 76 AA.
 XX AAAY9031:
 30-MAY-2000 (first entry)
 OS Homo sapiens.
 DE Amino acid sequence of chemokine receptor ligand MCP-2.
 XX
 KW Chemokine receptor; ligand; inflammatory response; immune effector cell;
 KW secondary tissue damage; central nervous system injury; MCP-2;
 KW CNS inflammatory disease; neurodegenerative disorder; heart disease;
 KW inflammatory eye disease; inflammatory bowel disease;
 KW inflammatory joint disease; inflammatory kidney; renal disease;
 KW inflammatory lung disease; inflammatory nasal disease;
 KW inflammatory thyroid disease; thyroiditis; cytokine-regulated cancer.
 XX
 OS Homo sapiens.
 PN W0200004926-A2.
 PD 04-FEB-2000.
 PF 21-JUL-1999: 99MO-CA00659.
 PR 22-JUL-1998: 98US-0120523.
 XX (USPTO) OSPREY PHARM LTD.
 XX

PI McDonald JR, Coggin PJ:
 XX WFI: 2000-182542/16.
 DR A new therapeutic agent comprising a conjugate for treating secondary
 XX tissue damage and other disease conditions like Alzheimer's disease,
 PT stroke, Parkinson's disease and atherosclerosis
 PI Disclosure: Page 60; 204pp; English.
 PS
 XX The present sequence represents a chemokine receptor ligand. The present
 CC ligand can be incorporated into the conjugates of the invention. The
 CC specification describes a conjugate, comprising a targeted agent and a
 CC chemokine receptor ligand. The conjugate binds to a chemokine receptor
 CC resulting in internalisation of the targeted agent in cells bearing the
 CC receptor. The conjugates are used for formulating a medicament or for
 CC treating disorders associated with inflammatory responses resulting from
 CC activation, proliferation and migration of immune effector cells. The
 CC disorders or disease states comprise secondary tissue damage such as
 CC central nervous system (CNS) injury, CNS inflammatory diseases,
 CC neurodegenerative disorders, heart disease, inflammatory eye diseases,
 CC inflammatory bowel diseases, inflammatory joint diseases, inflammatory
 CC kidney or renal diseases, inflammatory lung diseases, inflammatory
 CC nasal diseases, inflammatory thyroid disease such as thyroiditis, or
 CC cytokine-regulated cancers.
 XX
 SQ Sequence 76 AA:
 Query Match 100.0%; Score 25; DB 21; Length 76;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LEYSY 5
 DB 25 LEYSY 29
 RESULT 8
 AA021099
 ID AA021099 standard; Protein: 76 AA.
 XX AA021099:
 19-JUL-2002 (first entry),
 DE Protein of WT-MCP-2.
 XX
 KW RANTES; neuroprotective; anti-allergic; anti-inflammatory; anti-HIV; human;
 KW chemokine mutant; cationic site; multiple sclerosis; HIV infection;
 KW inflammatory disease; demyelinating disease; allergic.
 XX
 OS Homo sapiens.
 PN W0200228419-A2.
 PD 11-APR-2002.
 PF 03-OCT-2001: 2001WO-EP11426.
 PR 04-OCT-2000: 2000EP-0121995.
 XX (ISTE) ARS APPLIED RES SYSTEMS HOLDING NV.
 PA Proudfoot A, Wells TNC, Kosco-Villbois M:
 PI WPI: 2002-340073/37.
 DR A mutant of the human CC chemokine RANTES with two mutations in the
 PT cationic site of the 40% loop is used for treatment of multiple
 PT sclerosis and/or demyelinating diseases -
 XX Examples: Page 46; 46pp; English.
 XX

CC The invention relates to a truncated and mutated human RANTES (a CC
 CC chemokine mutant), comprising the amino sequence of 91 amino acids as
 CC given in the specification. The CC chemokine mutant RANTES, with two
 CC mutations in the cationic site is useful for the preparation of a
 CC pharmaceutical composition used in treating multiple sclerosis or other
 CC demyelinating diseases. The mutant with single mutations at cationic
 CC sites is used for the treatment of HIV infection and/or other allergic or
 CC inflammatory diseases. This sequence represents an example of a wild-type
 CC CC chemokine protein used in the invention.

XX Sequence 76 AA:

Query Match 100.0%; Score 25; DB 23; Length 76;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
 |||||
 DB 25 LESYT 29

RESULT 9

AA021109
 ID AA021109 standard; protein; 76 AA.

AC AA021109;

DT 19-JUL-2002 (first entry)

DE MCP-2 chemokine protein containing 40's loop and cationic site

KW RANTES: neuroprotective; anti-allergic; anti-inflammatory; anti-HIV; human;

KW chemokine mutant; cationic site; multiple sclerosis; HIV infection;

KW inflammatory disease; demyelinating disease; allergic

OS Homo sapiens.

Key Location/Qualifiers

Region 36..52
 /label= 40's loop

Misc-difference 46..49
 /label= Cationic site

W0200228419-A2.

11-APR-2002.

03-OCT-2001; 2001WO-EP11428.

04-OCT-2000; 2000EP-0141555.

(ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.

Proudfoot A, Wells TNC, Kosco-Vilbois M;

WPI: 2002-340073/37.

A mutant of the human CC chemokine RANTES with two mutations in the
 PT cationic site of the 40's loop is used for treatment of multiple
 PT sclerosis and/or demyelinating diseases -

PS Disclosure: Fig 1; 46pp; English.

CC The invention relates to a truncated and mutated human RANTES (a CC
 CC chemokine mutant), comprising the amino sequence of 91 amino acids as
 CC given in the specification. The CC chemokine mutant RANTES, with two
 CC mutations in the cationic site is useful for the preparation of a
 CC pharmaceutical composition used in treating multiple sclerosis or other
 CC demyelinating diseases. The mutant with single mutations at cationic
 CC sites is used for the treatment of HIV infection and/or other allergic or
 CC inflammatory diseases. This sequence represents an example of a CC
 CC chemokine protein containing a 40's loop and cationic site of the
 CC invention.

XX Sequence 76 AA:

Query Match 100.0%; Score 25; DB 23; Length 76;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
 |||||
 DB 25 LESYT 29

RESULT 10

AAG78390
 ID AAG78390 standard; protein; 76 AA.

AC AAG78390;

DT 30-MAY-2002 (first entry)

DE Human chemokine hMCP-2.

KW Antigen presenting cell; APC; chemotaxis; immune response; antigen;

KW chemokine; antitumor; cytostatic; antimicrobial; toxin;

KW pathogenic agent; cancer; melanoma; thyroid carcinoma;

KW renal cell carcinoma; tumor-associated antigen; human; hMCP-2.

OS Homo sapiens.

W020018082-A2.

01-NOV-2001.

20-APR-2001; 2001WO-US12000.

21-APR-2000; 2000US-1988 (wp).

12-APR-2001; 2001US-0834814.

(CHRM-) CHRM-CENTRIX INC.

Schall TJ, Talbot D;

WPI: 2002-034405/04.

Claim 7; Fig 2; 60pp; English.

CC The present sequence is that of a human chemokine polypeptide designated
 CC hMCP-2, used in the compositions of the invention. The specification
 CC describes the use of a composition containing an antigen presenting cell
 CC (APC)-chemotaxis for inducing an immune response to an antigen in a
 CC subject, in the manufacture of a medicament, where the APC-chemotaxis is
 CC a chemokine polypeptide or its variant, or a polynucleotide encoding the
 CC chemokine polypeptide or its variant. The invention has antitumor,
 CC cytostatic and antimicrobial activity. The compositions of the invention
 CC can contain one or more antigens for antigen containing polynucleotides
 CC and may be administered in the same mixture as the APC-chemotaxis or
 CC separately. Synthetic (inactive) or naturally occurring chemokines may be
 CC used. The polypeptides of the invention are used to induce an immune
 CC response by recruiting APCs to areas of antigen contact. The compositions
 CC of the invention are useful for providing protection from foreign
 CC infectious pathogenic agents (bacteria, virus, etc.) prior to expected
 CC or possible exposure; for treating cancers, melanomas,
 CC thyroid carcinomas, lung and breast cancers, renal cell carcinomas etc.,
 CC here the antigen present in the composition is a tumor associated
 CC antigen. The use of APC-chemotaxis provides an accelerated immune
 CC response in a host following administration of antigen, a more effective
 CC response to administration of exposure to very small quantities of
 CC an antigen (e.g., toxin or pathogen) due to increased antigen uptake by

CC APCs, and more effective anti-tumour therapies.
 XX
 SQ Sequence 76 AA:

Query Match 100.0%; Score 25; DB 23; Length 76;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LE5Y1 5
 11111
 Db 25 LE5Y1 29

RESULT 11
 AAG68353
 ID AAG68353 standard: peptide 76 AA
 XX
 AC AAG68353:
 VX 01-MAR-2002 (first entry)

Human monocyte chemotactic protein 2 (MCP2) amino acid sequence.

XX Human: mouse: monocyte chemotactic protein 2; hMCP2, MCP2, mC10, mMEC;
 KW chimeric chemokine; immune response; cytostatic; antitubercular;
 KW immunostimulant; vaccine; cancer; breast cancer; lung cancer;
 KW carcinoma; melanoma; tumour; hyaluronin-
 XX
 OS Homo sapiens.
 FN WO200160867-A2.
 PD 01-NOV-2001.
 XX
 FE 12-APR-2001; 2001WD-0512162.
 XX
 PR 21-APR-2000; 2000US-198839P.
 XX
 PA (CHEM-) CHEMOCENTRIX INC.
 PI Schall TJ, Talbot D;
 PI WPI: 2002-049246/06.

XX Inducing an immune response to an antigen to confer non-specific
 PT protection while the body is generating the adaptive response.
 PI Comprises administering a composition containing an antigen-presenting
 PI cell chemotaxin

Example 5; Fig 2; 60pp; English.

XX The present invention describes a method for inducing an immune response
 CC to an antigen in a subject. The method comprises administering a
 CC composition containing an antigen-presenting cell chemotaxin
 CC (APC-chemotaxin). Where the APC-chemotaxin is a chimeric polypeptide
 CC consisting of a polynucleotide encoding the chemokine polypeptide or
 CC its variant, the APC-chemotaxin has cytostatic, antimicrobial and
 CC immunostimulant activities, and can be used in vaccine production.
 CC The method can be used for inducing or enhancing an immune response,
 CC or for providing protection from exogenous foreign infectious pathogenic
 CC agents prior to expected or possible exposure. The method may be used to treat cancers
 CC displaying symptoms of exposure. The method may be used to treat cancers
 CC e.g. breast cancer, lung cancer, carcinomas, melanomas, and tumours.
 CC The present sequence represents a human monocyte chemotactic protein 2
 CC (hMCP2) amino acid sequence, which is used in an example from the
 CC present invention for the design of hybridines (chimeric chemokines).
 CC
 XX
 SQ Sequence 76 AA:

Query Match 100.0%; Score 25; DB 23; Length 76;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LE5Y1 5
 11111
 Db 25 LE5Y1 29

RESULT 12
 AAB15786
 ID AAB15786 standard: Protein 77 AA.
 XX
 AC AAB15786:
 XX
 DT 17-JAN-2001 (first entry)
 XX
 DE Human chemokine MCP-2 SEQ. ID NO: 17.

XX Macrophage recruitment; chemokine derivative; MCP-1; osteoporosis;
 XX monocyte chemoattractant protein-1; inflammation; atherosclerosis; HIV;
 KW AIDS; stroke; psoriasis; autoimmune disease; hypertension; endotoxaemia;
 KW basophil-mediated disease; myocardial infarction; acute ischaemia;
 KW rheumatoid arthritis; contraception.
 XX
 OS Homo sapiens.
 FN WO200042071-A2.
 PD 20-03L-2000.
 XX
 FE 12-JAN-2000; 2000WD-NSC008A1.
 XX
 PR 12-JAN-1999; 99US-0229071.
 PR 17-MAR-1999; 99US-0271172.
 PR 01-DEC-1999; 99US-0452416.
 XX
 PA (NEOR-) NEORX CORP.
 PI Granger DJ, Tatalick LM;
 PI WPI: 2000-499101/44.
 DR N-PSDB: AAA74886.

XX New peptide 3, amide and heterocyclic compounds and saccharide
 PT conjugates used for inhibiting chemokine induced activity and for
 PT treating e.g. stroke, vascular diseases, autoimmune diseases and tumour
 PT growth
 PI
 PI
 PS Example 1; Page 134; 387pp; English.

XX The present invention concerns the identification of a number of
 CC chemokines which can be used to produce derivatized, agonists and
 CC antagonists which are then useful in disease treatment. The chemokines
 CC include sequences AAB15786-815794, AAB15803-815813 and AAB15831-815845.
 CC These chemokine derivatives can be used to treat diseases such as
 CC autoimmune diseases, atherosclerosis, osteoporosis, HIV infection and
 CC AIDS, psoriasis, inflammatory diseases, hypertension, basophil-mediated
 CC diseases, endotoxaemia, myocardial infarction, acute ischaemia and
 CC rheumatoid arthritis, and can be used to prevent strokes and as
 CC contraceptives. The coding sequences for the chemokines can be used in
 CC gene therapy for the same diseases, as well as in the production of
 CC animal models.
 CC
 XX
 SQ Sequence 77 AA:

Query Match 100.0%; Score 25; DB 23; Length 77;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LE5Y1 5
 11111
 Db 26 LE5Y1 30

Search completed: April 8, 2003, 12:01:37
 Job time: 45.333 secs

Matches 5: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 QY 1 LE5YT 5
 |||||
 Db 26 LE5YT 30

RESULT 15

AAG68358

ID AAG68358 standard: peptide; 77 AA.

XX AAG68358:

DT 01-MAR-2002 (first entry)

DE Chimeric chemokine mMDC/1MCP2 amino acid sequence.

KW Human: mouse; monocytic chemotactic protein 2; 1MCP2; MCP2; mC10; mMDC;

KW chimeric chemokine; immune response; cytostatic; antimicrobial;

KW immunostimulant; vaccine; cancer; breast cancer; lung cancer;

KW carcinoma; melanoma; tumour; hybridoma.

XX Chimeric - Mus sp.

OS Chimeric - Homo sapiens.

XX WO200180887-A2.

PD 01-NOV-2001.

XX 12-APR-2001; 2001WO-US12162.

XX 21-APR-2000; 2000US-198939P.

XX (CHEM-) CHEMOCENTRIX INC.

XX Schall TJ; Talbot D;

XX WPI: 2002-049246/06.

XX Inducing an immune response to an antigen to confer non-specific

XX protection while the body is generating the adaptive response,

XX comprises administering a composition containing an antigen-presenting

XX cell chemotaxin -

XX Example 5: Fig 2; 60pp; English.

XX The present invention describes a method for inducing an immune response

XX to an antigen in a subject. The method comprises administering a

XX composition containing an antigen-presenting cell chemotaxin

XX (APC-chemotaxin), where the APC-chemotaxin is a chemokine polypeptide,

XX its variant, or a polynucleotide encoding the chemokine polypeptide or

XX its variant. The APC-chemotaxin has cytostatic, antimicrobial and

XX immunostimulant activities, and can be used in vaccine production.

XX The method can be used for inducing or enhancing an immune response,

XX or for providing protection from exogenous foreign infectious pathogenic

XX agents prior to expected or possible exposure, or to individuals

XX displaying symptoms of exposure. The method may be used to treat cancers

XX e.g. breast cancer, lung cancer, carcinomas, melanomas, and tumours.

XX The present sequence represents an exemplary chimeric chemokine

XX designated mMDC/1MCP2, which is given in an example from the present

XX invention for the design of hybridomas (chimeric chemokines).

SQ Sequence 77 AA:

Query Match 100.0%; Score 25; UB 23; Length 77;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 5: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LE5YT 5
|||||
Db 26 LE5YT 30

GenCorp Version 5.1.4-PS-457P
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein - protein search, using SW model

Run on: April 8 2003, 11:53:45 : Search time 15.467 seconds
(without alignments)
9.543 Million cell updates/sec

Title: US-09-647-749A-1

Perfect score: 25

Sequence: 1 LRSYT S

Scoring table: RUSOM62

Gapop 10.0 : Gapext 0.5

Searched: 262574 seqs, 29422222 residues

262574

ALIGNMENTS

Minimum DB seq length: 0

Maximum DB seq length: 400000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents AA:*

1: /cgn2_6/prodata/1/aa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/aa/5A_COMB.pep.*
3: /cgn2_6/prodata/1/aa/5A_COMB.pep.*
4: /cgn2_6/prodata/1/aa/5A_COMB.pep.*
5: /cgn2_6/prodata/1/aa/5A_COMB.pep.*
6: /cgn2_6/prodata/1/aa/5A_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	FP	ID	Description
1	25	100.0	74	2	US-08-615-232A-6
2	25	100.0	74	3	US-08-470-423-6
3	25	100.0	74	1	US-08-480-449-20
4	25	100.0	76	2	US-08-716-188-3
5	25	100.0	76	2	US-08-660-542-20
6	25	100.0	76	4	US-08-470-603-20
7	25	100.0	77	1	US-08-347-4928-9
8	25	100.0	77	2	US-08-421-144A-6
9	25	100.0	77	2	US-08-798-143-9
10	25	100.0	77	4	US-08-366-882A-12
11	25	100.0	255	4	US-09-327-857-488
12	25	100.0	301	2	US-08-951-871-6
13	25	100.0	634	4	US-08-134-001C-4612
14	25	100.0	2499	5	5182899-2
15	23	92.0	185	4	US-09-562-737-116
16	23	92.0	280	1	US-08-781-562-4
17	23	92.0	316	4	US-09-134-001C-3735
18	23	92.0	876	4	US-08-436-664-20
19	23	92.0	876	2	US-08-436-664-32
20	23	92.0	876	2	US-08-436-664-34
21	23	92.0	876	3	US-08-436-664-30
22	23	92.0	876	3	US-08-436-664-32
23	23	92.0	876	3	US-08-436-664-34
24	23	92.0	876	3	US-08-436-664-30
25	23	92.0	876	3	US-08-394-232A-32
26	23	92.0	876	3	US-08-394-232A-34
27	23	92.0	876	4	US-08-517-071-2

29	23	92.0	876	5	US-08-615-232A-6	Sequence 20, Appl
30	23	92.0	876	5	US-08-615-232A-6	Sequence 32, Appl
31	23	92.0	876	5	US-08-615-232A-6	Sequence 34, Appl
32	23	92.0	876	5	US-08-615-232A-6	Sequence 2, Appl
33	23	92.0	876	5	US-08-615-232A-6	Sequence 4, Appl
34	23	92.0	876	5	US-08-615-232A-6	Sequence 2, Appl
35	23	92.0	876	5	US-08-615-232A-6	Sequence 4, Appl
36	23	92.0	876	5	US-08-615-232A-6	Sequence 2, Appl
37	23	92.0	876	5	US-08-615-232A-6	Sequence 4, Appl
38	23	92.0	876	5	US-08-615-232A-6	Sequence 2, Appl
39	23	92.0	876	5	US-08-615-232A-6	Sequence 4, Appl
40	23	92.0	876	5	US-08-615-232A-6	Sequence 2, Appl
41	23	92.0	876	5	US-08-615-232A-6	Sequence 4, Appl
42	23	92.0	876	5	US-08-615-232A-6	Sequence 2, Appl
43	23	92.0	876	5	US-08-615-232A-6	Sequence 4, Appl
44	23	92.0	876	5	US-08-615-232A-6	Sequence 2, Appl
45	23	92.0	876	5	US-08-615-232A-6	Sequence 4, Appl

RESULT 1
US-08-615-232A-6
Sequence 6, Application US-08615232A
Patent No. 5993814
GENERAL INFORMATION:
APPLICANT: WILLIAMS, T. MOTHY J.
APPLICANT: GRIFFITHS-OLINSON, DAVID A.
APPLICANT: HEYMAN, JOHN I.
TITLE OF INVENTION: CHOLESTERIC CRYSTAL
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESS: RIXON & VANDERHAYE P. C.
STREET: 1100 NORTH GREE ROAD, 8TH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS-4.02
SOFTWARE: Patent In Release #1.0, Version #1.25 (HW)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 11-25-615-232A
FILING DATE: 13-AUG-1995
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08-931894
FILING DATE: 14-SEP-1993
APPLICATION NUMBER: 08-9408602
FILING DATE: 29-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 023955
REFERENCE/REGISTER NUMBER: 5993814
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-615-232A-6
Query Match: 100% Score 25, DB 23 Length 74:
Percent Similarity 100% Ident. No. 40:

Matches 5: Conservative 0: Mismatches 0: Indels 0: Caps 0:

QY 1 LESYT 5
|||||
DB 23 LESYT 27

RESULT 2

US-08-470-323-6
Sequence 6, Application US/08470323A
Patent No. 6031080
GENERAL INFORMATION:
APPLICANT: WILLIAMS, TIMOTHY J.
APPLICANT: JOSE, PETER J.
APPLICANT: GRIFFITHS-JOHNSON, DAVID A.
APPLICANT: HSUAN, JOHN J.
TITLE OF INVENTION: CHEMOTACTIC CYTOKINE
FILE REFERENCE: 550-33
CURRENT APPLICATION NUMBER: US-08-470-323A
CURRENT FILING DATE: 1995-06-06
EARLIER APPLICATION NUMBER: PCT/5894/02006
EARLIER FILING DATE: 1994-09-14
EARLIER APPLICATION NUMBER: GB 9318984.3
EARLIER FILING DATE: 1993-09-14
EARLIER APPLICATION NUMBER: GB 9406699.2
EARLIER FILING DATE: 1994-04-29
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 6
LENGTH: 74
TYPE: PRT
ORGANISM: human
US-08-470-323-6

Query Match 100.0%, Score 25; DB 3; Length 74;
Best Local Similarity 100.0%, Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
|||||
DB 23 LESYT 27

RESULT 3
US-08-480-449-20
Sequence 26, Application US/08480449
Patent No. 5686927
GENERAL INFORMATION:
APPLICANT: Godiska, Ronald
APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: MACROPHAGE DERIVED CHEMOKINE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,449
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE LOCAL NUMBER: 2786,72778
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: "Hu MCP-2"

US-08-480-449-20

Query Match 100.0%, Score 25; DB 1; Length 76;
Best Local Similarity 100.0%, Pred. No. 41;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
|||||
DB 25 LESYT 29

RESULT 4
US-08-716-188-3
Sequence 3, Application US/08716188
Patent No. 5908829
GENERAL INFORMATION:
APPLICANT: KELLY, RODNEY W.
TITLE OF INVENTION: USE OF MCP-1 FOR INDUCING RIPENING OF
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANIERHVE P.C.
STREET: 1100 NORTH GLENE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/716,188
FILING DATE: 30-SEP-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/5895/00733
FILING DATE: 31-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: G0 9406463.1
FILING DATE: 31-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 46,663
REFERENCE/LOCAL NUMBER: 117-219
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4091
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-716-188-3

Query Match 100.0%, Score 25; DB 2; Length 76;
Best Local Similarity 100.0%, Pred. No. 41;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
|||||
Db 25 LESYT 29

RESULT 5
US-08-660-542-20

Sequence 20, Affiliation US/08/660542
Patent No. 5932703

GENERAL INFORMATION:

APPLICANT: Godiska, Ronald

APPLICANT: Gray, Patrick W.

TITLE OF INVENTION: MACROPHAGE DERIVED CHEMOKINE AND CHEMOKINE

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borum

STREET: 6300 Sears Tower, 243 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/660,542

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/558,658

FILING DATE: 16-NOV-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/479,620

FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Gass, David A.

REGISTRATION NUMBER: 38,153

REFERENCE/DOCKET NUMBER: 27966-44219

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 76 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: misc:feature

OTHER INFORMATION: "Hu MCP-2"

US-08-660-542-20

QY 1 LESYT 5

|||||

Db 25 LESYT 29

RESULT 6

US-08-479-603-20

Sequence 20, Application US/08479603

Patent No. 6320023

GENERAL INFORMATION:

APPLICANT: Godiska, Ronald

APPLICANT: Gray, Patrick W.

TITLE OF INVENTION: MACROPHAGE DERIVED CHEMOKINE AND CHEMOKINE

QY 1 LESYT 5
|||||
Db 25 LESYT 29

RESULT 7
US-08-347-492B-9

Sequence 9, Affiliation US/0847492B
Patent No. 5602008

GENERAL INFORMATION:

APPLICANT: Wilde, Craig J.

APPLICANT: Hawkins, Phillip F.

APPLICANT: Bandman, Olga

APPLICANT: Sellman, Jeffrey J.

TITLE OF INVENTION: EXPRESSED CHEMOKINES, THEIR

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESS: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: U.S.

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Vers. 1.5

CURRENT APPLICATION DATA:

QY 1 LESYT 5
|||||
Db 25 LESYT 29

Query Match 100.0%; Score 25; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-08-479-603-20

QY 1 LESYT 5
|||||
Db 25 LESYT 29

RESULT 5
US-08-660-542-20

Sequence 20, Affiliation US/08660542
Patent No. 5932703

GENERAL INFORMATION:

APPLICANT: Godiska, Ronald

APPLICANT: Gray, Patrick W.

TITLE OF INVENTION: MACROPHAGE DERIVED CHEMOKINE AND CHEMOKINE

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borum

STREET: 6300 Sears Tower, 243 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/660,542

FILING DATE:

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Gass, David A.

REGISTRATION NUMBER: 38,153

REFERENCE/DOCKET NUMBER: 27966/32780

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 76 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: misc:feature

OTHER INFORMATION: "Hu MCP-2"

APPLICATION NUMBER: US/08/747 492B
FILING DATE: 29-NOV-1994
PRIOR APPLICATION DATA: 08/421 144A-6
APPLICATION NUMBER: 08/421 241
FILING DATE: 07-SEP-1994
APPLICATION NUMBER: 08/420 011
FILING DATE: 05-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Luther, Barbara J.
REGISTRATION NUMBER: 33,954
REFERENCE/DOCKET NUMBER: PF-0024
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-852-0195
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: GI 126829
US-08-447-492B-9

Query Match
Best Local Similarity 100.0%; Score 25; DB 1; Length 77;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEAST 5
DB 26 LEAST 30

RESULT 8
US-08-421-144A-6
Sequence 6, Application US/08421144A
Patent No. 5874211
GENERAL INFORMATION:
APPLICANT: BANDMAN, OLGA
APPLICANT: COLEMAN, ROGER
APPLICANT: STUART, SUSAN G.
TITLE OF INVENTION: NEW CHEMOKINE EXPRESSED IN EOSINOPHILS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/421,144A
FILING DATE: 13-APR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Luther, Barbara J.
REGISTRATION NUMBER: 33954
REFERENCE/DOCKET NUMBER: PF-0031 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-852-0195
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-421-144A-6

Query Match
Best Local Similarity 100.0%; Score 25; DB 2; Length 77;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEAST 5
DB 26 LEAST 30

RESULT 9
US-08-798-143-9
Sequence 9, Application US/08798143
Patent No. 5936068
GENERAL INFORMATION:
APPLICANT: Wilde, Craig J.
APPLICANT: Hawkins, Phillip R.
APPLICANT: Bandman, Olga
APPLICANT: Seilhamer, Jeffrey J.
TITLE OF INVENTION: EXPRESSED CHEMOKINES, THEIR
TITLE OF INVENTION: PROMOTION AND USES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/798,143
FILING DATE: 10-FEB-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/447,492
FILING DATE: 29-NOV-1994
APPLICATION NUMBER: 08/403,241
FILING DATE: 07-SEP-1994
APPLICATION NUMBER: 08/420,011
FILING DATE: 05-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Luther, Barbara J.
REGISTRATION NUMBER: 33,954
REFERENCE/DOCKET NUMBER: PF-0024
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-852-0195
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: GI 126829
US-08-798-143-9

Query Match
Best Local Similarity 100.0%; Score 25; DB 2; Length 77;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEAST 5
DB 26 LEAST 30

```
Db 26 LEAST 30

RESULT 10
US-09-366-487A-12
Sequence 12, Application US/09366487A
Patent No. 6403782
GENERAL INFORMATION:
APPLICANT: LUSTER, ANDREW D.
APPLICANT: LEDER, PHILIP
APPLICANT: POTENBERG, MARC
APPLICANT: GARCIA, EDUARDO
TITLE OF INVENTION: EOTAXIN: AN EOSINOPHIL CHEMOKINE/TILANT
FILE REFERENCE: 00483/025002
CURRENT APPLICATION NUMBER: US/09/366-487A
PRIOR FILING DATE: 1999-08-04
PRIOR APPLICATION NUMBER: 60/090,444
PRIOR FILING DATE: 1995-06-22
PRIOR APPLICATION NUMBER: 08/522,713
PRIOR FILING DATE: 1995-09-01
PRIOR APPLICATION NUMBER: 08/522,714
PRIOR FILING DATE: 1998-06-16
NUMBER OF SEQ ID NOS: 27
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 13
LENGTH: 77
TYPE: PRT
ORGANISM: Homo sapiens
US-09-366-487A-12

Query Match 100.0% Score 25; DB 4; Length 77;
Best local similarity 100.0% Freq No. 42;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEAST 5
|||||
Db 26 LEAST 30

RESULT 11
US-09-227-457-488
Sequence 488, Application US/09227457
Patent No. 6342591
GENERAL INFORMATION:
APPLICANT: FISCHER et al.
TITLE OF INVENTION: 123 Human Secreted Proteins
FILE REFERENCE: P2010P1
CURRENT APPLICATION NUMBER: US/09/227,357
CURRENT FILING DATE: 1999-01-08
EARLIER APPLICATION NUMBER: PCT/US98/13684
EARLIER FILING DATE: 1998-07-07
EARLIER APPLICATION NUMBER: 60/071,926
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,793
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,925
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,929
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,932
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,916
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,940
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08

Db 163 LEAST 167

RESULT 12
US-08-951-871-6
Sequence 5, Application US/08951871
Patent No. 5866398
GENERAL INFORMATION:
APPLICANT: XIAO, SHUANG-WANG
APPLICANT: XIAO, JIAN-PING
TITLE OF INVENTION: METHOD FOR CLONING
TITLE OF INVENTION: AND PRODUCING THE 1231 RESTRICTION ENDOGENE IN F.
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESSES:
ADDRESSEE: New England Biolabs, Inc.
STREET: 32 Tozer Road
```

```

CITY: Beverly
STATE: MA
COUNTRY: US
ZIP: 01915
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/951.871
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Gregory D
REGISTRATION NUMBER: 30901
REFERENCE/DOCKET NUMBER: NEB-131
TELECOMMUNICATION INFORMATION:
TELEPHONE: 978-927-5054
TELEFAX: 978-927-1705
TELEX:
INFORMATION FOR SEQ. ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 301 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-951-871-6

Query Match          100.0%  Score 25.  DB 2:  Length 301:
Best Local Similarity 100.0%  Pred. No. 1.7e+02:
Matches 5:  Conservative 0:  Mismatches 0:  Indels 0:  Gaps 0:

QY 1 LESTY 5
DB 216 LESTY 220

RESULT 13
US-09-134-001C-4633
Sequence 4633, Application US/09/134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSIS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIORITY APPLICATION NUMBER: US 60/7064,964
PRIORITY FILING DATE: 1997-11-08
PRIORITY APPLICATION NUMBER: US 60/7055,779
PRIORITY FILING DATE: 1997-08-14
NUMBER OF SEQ. ID NOS: 5674
SEQ. ID NO 4633
LENGTH: 538
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4633

Query Match          100.0%  Score 25:  DB 4:  Length 538:
Best Local Similarity 100.0%  Pred. No. 3e+02:
Matches 5:  Conservative 0:  Mismatches 0:  Indels 0:  Gaps 0:

QY 1 LESTY 5
DB 291 LESTY 295

```

```

RESULT 14
5180808-2
Patent No. 5180808
APPLICANT: ROOSLAHNT, EREKI L.
TITLE OF INVENTION: VESICULAR CORE PROTEIN, NUCLEIC ACID
SEQUENCES ENCODING THE SAME, NUCLEIC ACID PROBES, ANTI-VESSICAN
ANTIBODIES, AND METHODS OF DETECTING THE SAME
NUMBER OF SEQUENCES: 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/441,179
FILING DATE: 27-NOV-1989
SEQ. ID NO: 2
LENGTH: 2409
5180808-2

Query Match          100.0%  Score 25:  DB 6:  Length 2409:
Best Local Similarity 100.0%  Pred. No. 1.4e+03:
Matches 5:  Conservative 0:  Mismatches 0:  Indels 0:  Gaps 0:

QY 1 LESTY 5
DB 1203 LESTY 1207

RESULT 15
US-09-562-737-116
Sequence 116, Application US/09562737
Patent No. 6428967
GENERAL INFORMATION:
APPLICANT: Herz, Joachim
TITLE OF INVENTION: LDL Receptor Signaling Pathways
FILE REFERENCE: UTSW0708
CURRENT APPLICATION NUMBER: US/09/562,737
CURRENT FILING DATE: 2000-05-01
NUMBER OF SEQ. ID NOS: 132
SOFTWARE: PatentIn Ver. 2.1
SEQ. ID NO 116
LENGTH: 185
TYPE: PRT
ORGANISM: Artificial Sequence
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-562-737-116

Query Match          92.7%  Score 23:  DB 4:  Length 185:
Best Local Similarity 80.1%  Pred. No. 3e+02:
Matches 4:  Conservative 1:  Mismatches 0:  Indels 0:  Gaps 0:

QY 1 LESTY 5
DB 86 MESTY 90

Search completed: April- 8, 2003, 12:05:51
Job time : 16.4167 secs

```

Genome version 5.1.4.25.4578
Copyright (c) 1993 - 2000. Genengen Ltd.

OM protein - protein search, using sw model

Run on: April 8 2003 12:03:50, Search time 9.10667 seconds
(without alignments) 33.347 Million coll updates/sec

Title: US-09-647-749a-1

Perfect score: 25

Sequence: 1 LESYT 5

Scoring table:

Gap: 10.0, Offset 0.0

Searched: 248912 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248912

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications-AA:

1: /cgn2_6/pdata/1/pubpa/US08_NEW_PUB pep.*
2: /cgn2_6/pdata/1/pubpa/US08_NEW_PUB pep.*
3: /cgn2_6/pdata/1/pubpa/US08_NEW_PUB pep.*
4: /cgn2_6/pdata/1/pubpa/US08_NEW_PUB pep.*
5: /cgn2_6/pdata/1/pubpa/US08_NEW_PUB pep.*
6: /cgn2_6/pdata/1/pubpa/US08_NEW_PUB pep.*
7: /cgn2_6/pdata/1/pubpa/US08_NEW_PUB pep.*
8: /cgn2_6/pdata/1/pubpa/US08_NEW_PUB pep.*
9: /cgn2_6/pdata/1/pubpa/US08_NEW_PUB pep.*
10: /cgn2_6/pdata/1/pubpa/US08_NEW_PUB pep.*
11: /cgn2_6/pdata/1/pubpa/US08_NEW_PUB pep.*
12: /cgn2_6/pdata/1/pubpa/US08_NEW_PUB pep.*
13: /cgn2_6/pdata/1/pubpa/US08_NEW_PUB pep.*
14: /cgn2_6/pdata/1/pubpa/US08_NEW_PUB pep.*

Free N is the number of results predicted by query to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	74	10	US-09-647-741-45720
2	25	100.0	74	10	US-09-647-741-45720
3	25	100.0	74	10	US-09-647-741-45720
4	25	100.0	74	10	US-09-647-741-45720
5	25	100.0	74	10	US-09-647-741-45720
6	25	100.0	74	10	US-09-647-741-45720
7	25	100.0	74	10	US-09-647-741-45720
8	25	100.0	74	10	US-09-647-741-45720
9	25	100.0	74	10	US-09-647-741-45720
10	25	100.0	74	10	US-09-647-741-45720
11	25	100.0	74	10	US-09-647-741-45720
12	25	100.0	74	10	US-09-647-741-45720
13	25	100.0	74	10	US-09-647-741-45720
14	25	100.0	74	10	US-09-647-741-45720
15	25	100.0	74	10	US-09-647-741-45720
16	25	100.0	74	10	US-09-647-741-45720
17	25	100.0	74	10	US-09-647-741-45720
18	25	100.0	74	10	US-09-647-741-45720
19	25	100.0	74	10	US-09-647-741-45720

ALIGNMENTS

20	23	92.0	486	10	US-09-647-741-45720	Sequence 598, App
21	23	92.0	486	10	US-09-647-741-45720	Sequence 10668, A
22	23	88.0	12	4	US-09-647-741-45720	Sequence 40, App1
23	23	88.0	12	4	US-09-647-741-45720	Sequence 2, App11
24	23	88.0	12	4	US-09-647-741-45720	Sequence 2, App11
25	23	88.0	12	4	US-09-647-741-45720	Sequence 19, App1
26	23	88.0	12	4	US-09-647-741-45720	Sequence 4547, A
27	23	88.0	12	4	US-09-647-741-45720	Sequence 41, App1
28	23	88.0	12	4	US-09-647-741-45720	Sequence 45840, A
29	23	88.0	12	4	US-09-647-741-45720	Sequence 10, App1
30	23	88.0	12	4	US-09-647-741-45720	Sequence 147, App
31	23	88.0	12	4	US-09-647-741-45720	Sequence 708, App
32	23	88.0	12	4	US-09-647-741-45720	Sequence 225, App
33	23	88.0	12	4	US-09-647-741-45720	Sequence 225, App
34	23	88.0	12	4	US-09-647-741-45720	Sequence 225, App
35	23	88.0	12	4	US-09-647-741-45720	Sequence 225, App
36	23	88.0	12	4	US-09-647-741-45720	Sequence 225, App
37	23	88.0	12	4	US-09-647-741-45720	Sequence 225, App
38	23	88.0	12	4	US-09-647-741-45720	Sequence 225, App
39	23	88.0	12	4	US-09-647-741-45720	Sequence 225, App
40	23	88.0	12	4	US-09-647-741-45720	Sequence 225, App
41	23	88.0	12	4	US-09-647-741-45720	Sequence 225, App
42	23	88.0	12	4	US-09-647-741-45720	Sequence 225, App
43	23	88.0	12	4	US-09-647-741-45720	Sequence 225, App
44	23	88.0	12	4	US-09-647-741-45720	Sequence 225, App
45	23	88.0	12	4	US-09-647-741-45720	Sequence 225, App

```

PRIOR APPLICATION NUMBER: P01/US01/000670
PRIOR FILING DATE: 2001-01-20
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-10
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 46020
LENGTH: 68
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC008494.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
OTHER INFORMATION: EST_HUMAN HIT: AA57020.1, EVALUATE 1.00e-07
+09-864-761-46020

```

```

Query Match          100.0% Score 25; DB 10; Length 68;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LESYT 5
DB 25 LESYT 29

```

```

RESULT 2
US-09-195-457-6
Sequence 6; Application US/09195457
GENERAL INFORMATION:
APPLICANT: WILLIAMS, TIMOTHY J.
APPLICANT: JOSE, PETER J.
APPLICANT: GRIFFITHS-JOHNSON, DAVID A.
APPLICANT: HSUAN, JOHN J.
TITLE OF INVENTION: CHEMOTACTIC CYTOKINE
FILE REFERENCE: 550-33
CURRENT APPLICATION NUMBER: US/09/195,457
CURRENT FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 08/470,323
PRIOR FILING DATE: 1995-06-06
PRIOR APPLICATION NUMBER: P7/95/470,323
PRIOR FILING DATE: 1994-09-14
PRIOR APPLICATION NUMBER: GB 9318984.3
PRIOR FILING DATE: 1993-09-14
PRIOR APPLICATION NUMBER: GB 94086902.2
PRIOR FILING DATE: 1994-04-29
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 6
LENGTH: 74
TYPE: PRT
ORGANISM: human
US-09-195-457-6

```

```

Query Match          100.0% Score 25; DB 10; Length 74;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LESYT 5
DB 23 LESYT 27

```

```

RESULT 3
US-09-792-793A-21
Sequence 21; Application US/09792793A
Patent No. US20020168170A1
GENERAL INFORMATION:
APPLICANT: McDonald, John R.
APPLICANT: Coggin, Philip

```

```

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE
FILE REFERENCE: 25020-601D
CURRENT APPLICATION NUMBER: US/09/792,793A
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 93
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 21
LENGTH: 76
TYPE: PRT
ORGANISM: homo sapien
FEATURE:
OTHER INFORMATION: Human Chemokine Polypeptide; MCP-2
US-09-792-793A-21

```

```

Query Match          100.0% Score 25; DB 9; Length 76;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LESYT 5
DB 25 LESYT 29

```

```

RESULT 4
US-08-927-939-17
Sequence 17; Application US/08927939
Patent No. US2001000640A1
GENERAL INFORMATION:
APPLICANT: Granger, David L.
APPLICANT: Tatalock, Lauren Marie
TITLE OF INVENTION: Compounds and methods to inhibit or
FILE REFERENCE: 295-022051
CURRENT APPLICATION NUMBER: US/08/927,939
CURRENT FILING DATE: 1997-06-11
NUMBER OF SEQ ID NOS: 83
SOFTWARE: FASTED for Windows Version 3.0
SEQ ID NO 17
LENGTH: 77
TYPE: PRT
ORGANISM: Homo sapiens
US-08-927-939-17

```

```

Query Match          100.0% Score 25; DB 9; Length 77;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LESYT 5
DB 26 LESYT 30

```

```

RESULT 5
US-10-057-275-10
Sequence 10; Application US/10057275
Patent No. US2002015545A1
GENERAL INFORMATION:
APPLICANT: Coleman, Roger
APPLICANT: Bandman, Olga
APPLICANT: Wilde, Craig G.
TITLE OF INVENTION: NEW CHEMOKINES EXPRESSED IN PANCREAS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Poulter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

```


OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057,275
FILING DATE: 25-Jan-2002
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/390,740A
FILING DATE: February 17, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Luther, Barbara J.
REGISTRATION NUMBER: 33,954
REFERENCE/DOCKET NUMBER: FF-0027 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-852-0195
TELEFAX: 415-852-0195
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: MCP-2
SEQUENCE DESCRIPTION: SEQ ID NO: 10.

Query Match 100.0% Score 25; DB 9; Length 77;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
DB 26 LESYT 30

RESULT 6
US-09-834-794A-24
Sequence 24, Application US/09/044,794A
Publication No. US2003026777A1
GENERAL INFORMATION:
APPLICANT: Lawrence, Papsidero
APPLICANT: Lyn, Dyster
TITLE OF INVENTION: Detection and Treatment of Breast Cancer
FILE REFERENCE: 3380/11127-US4
CURRENT APPLICATION NUMBER: US/09/044,794A
CURRENT FILING DATE: 2001-04-13
PRIORITY APPLICATION NUMBER: 09/146,580
PRIORITY FILING DATE: 1998-09-03
PRIORITY APPLICATION NUMBER: 60/071,899
PRIORITY FILING DATE: 1998-01-20
PRIORITY APPLICATION NUMBER: 60/092,155
PRIORITY FILING DATE: 1998-07-09
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.0
SEQ ID NO 24
LENGTH: 77
TYPE: PRT
ORGANISM: Homo sapiens
US-09-834-794A-24

Query Match 100.0% Score 25; DB 9; Length 77;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
DB 26 LESYT 30

RESULT 7

US-09-920-137A-10
Sequence 10, Application US/09/020137A
Publication No. US2003094977A1
GENERAL INFORMATION:
APPLICANT: Coleman, Roger
APPLICANT: Bandman, Olga
APPLICANT: Wilde, Craig
TITLE OF INVENTION: NEW PHARMACINES EXPRESSED IN HUMAN
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Icyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,137A
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Luther, Barbara J.
REGISTRATION NUMBER: 33,954
REFERENCE/DOCKET NUMBER: FF-0027 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-852-0195
TELEFAX: 415-852-0195
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: MCP-2
US-09-920-137A-10

Query Match 100.0% Score 25; DB 9; Length 77;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
DB 26 LESYT 30

RESULT 8
US-09-834-795A-24
Sequence 24, Application US/09/044,795A
Patent No. US2002007670A1
GENERAL INFORMATION:
APPLICANT: Lawrence, Papsidero
APPLICANT: Lyn, Dyster
TITLE OF INVENTION: Detection and Treatment of Breast Cancer
FILE REFERENCE: 3380/11127-US4
CURRENT APPLICATION NUMBER: US/09/044,795A
CURRENT FILING DATE: 2001-04-12
PRIORITY APPLICATION NUMBER: 09/146,580
PRIORITY FILING DATE: 1998-09-03
PRIORITY APPLICATION NUMBER: 60/071,899
PRIORITY FILING DATE: 1998-01-20
PRIORITY APPLICATION NUMBER: 60/092,155
PRIORITY FILING DATE: 1998-07-09
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.0
SEQ ID NO 24

us-09-647-749a-1.rapb

Score	DB	Length
100.08	25	77
100.08	Pred. No. 39	
100.08		Indels

W	1	TEST	5
Ed	26	TEST	30

```

Query Match      100.0%  Score 25;  DB 9;  Length 109;
similarity      100.0%  Pval No. 55;
Best local      0;
Matches         5;  Conservative 0;  Indels 0;  Gaps 0

```

US-09-764-R48-875
Sequence 875, Application US/09764868
Patent No. US2002016871A1
GENERAL INFORMATION:

```

1 FILE REFERENCE: PT232
2 CURRENT APPLICATION NUMBER: US/09/764,868
3 CURRENT FILING DATE: 2001-01-17
4 Prior application data removed - refer to PALM or file wrapper
5 NUMBER OF SEQ ID NOS: 1510
6 SOFTWARE: PatentIn Ver. 2.0
7 SEQ ID NO: 875
8 LENGTH: 151
9 TYPE: PRT
10 ORGANISM: Homo sapiens
11 FEATURE:
12 NAME/KEY: SITE
13 LOCATION: (117)
14 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
15 NAME/KEY: SITE
16 LOCATION: (127)
17 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
18
19 US-09-764,868-875

```

QY	1	LESYT	5
Db	136	LESYT	140

RESULT 11
US-09-791-279-201
: Sequence 201, Application US/00/11-79
: Publication No. US20030050456A1

```

APPLICANT: 14401-10101
AFFILIANT: 14401-10101
TITLE OF INVENTION: NO. US2004005456A1 3 Protein-coupled receptors
FILE REFERENCE: 00048-051
CURRENT APPLICATION NUMBER: 000000791, 279
CURRENT FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 000144, 715
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 000144725
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 000184, 712
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 000184, 602
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 000184, 604
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 000184, 822
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 000184, 710
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 000184, 689
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 000184, 690
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 000184, 716
PRIOR FILING DATE: 2000-02-24
NUMBER OF SEQ ID NOS: 220
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 201
LENGTH: 206
TYPE: PRT
ORGANISM: Homo sapiens
US-09-791-279-201

```

QY	1	LESYT	5
		11111	
Db	155	LESYT	159

```

RESULT 12
US-09-895-913A-184
: Sequence 184: Application 09/04895913A
: Patent No.: US20020160456A1
: GENERAL INFORMATION:
: APPLICANT: Kleanthous, Harold
: APPLICANT: Al-Garawi, Amal
: APPLICANT: Miller, Charles
: APPLICANT: Tomb, Jean Francois
: APPLICANT: Oomen, Raymond P.
: TITLE OF INVENTION: Identification of Polynucleotides

```

TITLE OF INVENTION: Encoding No. US20020196456A1 Helicobacter Polypeptides in the
 FILE REFERENCE: 0514/045002
 CURRENT APPLICATION NUMBER: US/09/895,913A
 CURRENT FILING DATE: 2001-06-29
 PRIOR APPLICATION NUMBER: US 08/781,227
 PRIOR FILING DATE: 1997-06-24
 NUMBER OF SEQ ID NOS: 368
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO: 184
 LENGTH: 249
 TYPE: PRT
 ORGANISM: Helicobacter pylori
 US-09-895-913A-184

Query Match 100.0% Score 25; DB 9; Length 249;
 Best Local Similarity 100.0% Freq. No. 1 25002;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 LEAST 5
 11111
 b 204 LEAST 208

RESULT 13
 US-09-983-802-488
 Sequence 488 Application US/0983802
 Publication No. US20020196456A1
 GENERAL INFORMATION:
 APPLICANT: Fischer et al.
 TITLE OF INVENTION: 123 Human Secreted Proteins
 FILE REFERENCE: P2101P1
 CURRENT APPLICATION NUMBER: US-09/983,802
 CURRENT FILING DATE: 2001-10-25
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/227,357
 PRIOR FILING DATE: EARLIER FILING DATE: 1998-01-08
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/US98/13684
 PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-07
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,926
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,793
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,925
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,920
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,918
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,920
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,733
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,795
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,919
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,928
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,722
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,723

PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,948
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,949
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,954
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,950
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,947
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,964
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/056,360
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,984
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/056,374
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/056,785
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/056,664
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/056,660
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/056,661
 NUMBER OF SEQ ID NOS: 672
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 488
 LENGTH: 255
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-983-802-488

Query Match 100.0% Score 25; DB 9; Length 255;
 Best Local Similarity 100.0% Freq. No. 1 38702;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LEAST 5
 11111
 Db 163 LEAST 167

RESULT 14
 US-09-738-626-6675
 Sequence 6675 Application US/09738675
 Publication No. US20020197604A1
 GENERAL INFORMATION:
 APPLICANT: NAKAYAMA, SATOSHI
 APPLICANT: MIZOGUCHI, HIRASHI
 APPLICANT: ANDO, SEIKO
 APPLICANT: HAYASHI, MIKIKO
 APPLICANT: OCHIDA, KEIKO
 APPLICANT: YOKOI, HARUHIKO
 APPLICANT: TATEISHI, NAOKO
 APPLICANT: SENOH, AKIHITO
 APPLICANT: IKEDA, MASATO
 APPLICANT: OZAKI, AKIO
 TITLE OF INVENTION: NOVEL POLYPEPTIDES
 FILE REFERENCE: 249-125
 CURRENT APPLICATION NUMBER: 09/738,675
 CURRENT FILING DATE: 2000-11-18
 PRIOR APPLICATION NUMBER: JP 99/377484
 PRIOR FILING DATE: 1999-12-14
 PRIOR APPLICATION NUMBER: JP 00/159162
 PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: JP 00/159162
 PRIOR FILING DATE: 2000-08-11
 NUMBER OF SEQ ID NOS: 7059
 SOFTWARE: PatentIn ver. 3.0

Tue Apr 8 16:01:57 2003

us-09-647-749a-1.rapb

SEQ ID NO 6675
LENGTH: 332
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6675

Query Match 100.0% Score 25; DB 9; Length 332;
Best Local Similarity 100.0%; Pctd. No. 17e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
11111
DB 234 LESYT 238

RESULT 15
US-10-160-501-14
Sequence 14, Application US/10160501
Publication No. US20030059919A1

GENERAL INFORMATION:

APPLICANT: Meyers, Rachel E.
APPLICANT: Rudolph-Owen, Laura A.
TITLE OF INVENTION: NOVEL HUMAN 39228, 21956, 25856, 22244, 8701, 32263,
TITLE OF INVENTION: 50250, 55158, 47765, 62088, 50566, AND 48118
TITLE OF INVENTION: MOLECULES AND USES THEREFOR

FILE REFERENCE: MNI-250

CURRENT APPLICATION NUMBER: US/10/160, 501

CURRENT FILING DATE: 2002-05-30

PRIOR APPLICATION NUMBER: US 09/838, 573

PRIOR FILING DATE: 2002-04-18

PRIOR APPLICATION NUMBER: US 60/197, 747

PRIOR FILING DATE: 2000-04-18

PRIOR APPLICATION NUMBER: US 09/870, 133

PRIOR FILING DATE: 2001-05-29

PRIOR APPLICATION NUMBER: US 60/207, 649

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/870, 130

PRIOR FILING DATE: 2001-05-29

PRIOR APPLICATION NUMBER: US 60/207, 640

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/862, 535

PRIOR FILING DATE: 2001-05-21

PRIOR APPLICATION NUMBER: US 60/205, 961

PRIOR FILING DATE: 2000-05-19

PRIOR APPLICATION NUMBER: US 09/870, 383

PRIOR FILING DATE: 2001-05-29

PRIOR APPLICATION NUMBER: US 60/207, 506

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/860, 821

PRIOR FILING DATE: 2001-05-18

PRIOR APPLICATION NUMBER: US 60/205, 449

PRIOR FILING DATE: 2000-05-19

PRIOR APPLICATION NUMBER: US 09/870, 110

PRIOR FILING DATE: 2001-05-29

PRIOR APPLICATION NUMBER: US 60/207, 650

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/907, 509

PRIOR FILING DATE: 2001-07-16

PRIOR APPLICATION NUMBER: US 60/218, 385

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: US 09/945, 327

PRIOR FILING DATE: 2001-08-31

PRIOR APPLICATION NUMBER: US 60/229, 425

PRIOR FILING DATE: 2000-08-31

PRIOR APPLICATION NUMBER: US 60/318, 581

PRIOR FILING DATE: 2001-09-10

NUMBER OF SEQ ID NOS: 41

SOFTWARE: FastSeq Version 4.0

SEQ ID NO 14

LENGTH: 467

TYPE: PRT

ORGANISM: Homo sapiens

US-10-160-501-14

Query Match 100.0% Score 25; DB 9; Length 467;
Best Local Similarity 100.0%; Pctd. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
11111
DB 248 LESYT 252

Search completed: April 8, 2004 12:06:20
Job time: 10.1667 secs

Genome version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

1 protein - protein search using SW model

on: April 8, 2003, 11:59:49 : Search time 21.5833 seconds

(without alignments)
9.544 Million cell updates/sec

Sequence: 1 KEYPTS 7

oring table: BL/SCMG2
Gapop 10.0 : Gapext 0.5

arched: 262574 seqs, 29422922 residues

nal number of hits satisfying chosen parameters: 262574

imum DB seq length: 0
imum DB seq length: 2000000000

st-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

2: 100% 100% 100% 100% 100% 100% 100% 100% 100% 100%
3: 100% 100% 100% 100% 100% 100% 100% 100% 100% 100%
4: 100% 100% 100% 100% 100% 100% 100% 100% 100% 100%
5: 100% 100% 100% 100% 100% 100% 100% 100% 100% 100%

Fold: No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	88.9	204	US-09-134-001C-3833	Sequence 3833, Ap
2	29	85.6	703	US-09-134-001C-3015	Sequence 3015, Ap
3	28	77.8	210	US-08-856-841-9	Sequence 9, Appl
4	28	77.8	218	US-08-811-583-3	Sequence 3, Appl
5	28	77.8	227	US-08-856-841-15	Sequence 15, Appl
6	28	77.8	250	US-08-856-841-12	Sequence 12, Appl
7	28	77.8	276	US-08-935-263-16	Sequence 16, Appl
8	28	77.8	276	US-09-594-185-16	Sequence 16, Appl
9	28	77.8	311	US-08-605-284B-18	Sequence 18, Appl
10	28	77.8	387	US-08-856-841-17	Sequence 17, Appl
11	28	77.8	486	US-08-856-841-19	Sequence 19, Appl
12	28	77.8	500	US-08-856-841-18	Sequence 18, Appl
13	28	77.8	501	US-08-856-841-18	Sequence 18, Appl
14	28	77.8	584	US-09-413-452-2	Sequence 2, Appl
15	28	77.8	584	US-09-413-452-2	Sequence 2, Appl
16	28	77.8	584	US-09-413-452-2	Sequence 2, Appl
17	28	77.8	584	US-09-413-452-2	Sequence 2, Appl
18	28	77.8	584	US-09-413-452-2	Sequence 2, Appl
19	28	77.8	584	US-09-413-452-2	Sequence 2, Appl
20	28	77.8	584	US-09-413-452-2	Sequence 2, Appl
21	28	77.8	584	US-09-413-452-2	Sequence 2, Appl
22	28	77.8	584	US-09-413-452-2	Sequence 2, Appl
23	28	77.8	584	US-09-413-452-2	Sequence 2, Appl
24	28	77.8	584	US-09-413-452-2	Sequence 2, Appl
25	28	77.8	584	US-09-413-452-2	Sequence 2, Appl
26	28	77.8	584	US-09-413-452-2	Sequence 2, Appl
27	28	77.8	584	US-09-413-452-2	Sequence 2, Appl

28	27	75.0	62	4	US-09-134-001C-3833	Sequence 3833, Ap
29	27	75.0	66	1	US-08-856-841-9	Sequence 9, Appl
30	27	75.0	66	1	US-08-856-841-9	Sequence 9, Appl
31	27	75.0	67	4	US-09-134-001C-3833	Sequence 3833, Ap
32	27	75.0	67	4	US-09-134-001C-3833	Sequence 3833, Ap
33	27	75.0	68	2	US-08-856-841-18	Sequence 18, Appl
34	27	75.0	68	2	US-08-856-841-18	Sequence 18, Appl
35	27	75.0	68	2	US-08-856-841-18	Sequence 18, Appl
36	27	75.0	68	2	US-08-856-841-18	Sequence 18, Appl
37	27	75.0	68	2	US-08-856-841-18	Sequence 18, Appl
38	27	75.0	68	2	US-08-856-841-18	Sequence 18, Appl
39	27	75.0	68	2	US-08-856-841-18	Sequence 18, Appl
40	27	75.0	69	4	US-08-856-841-18	Sequence 18, Appl
41	27	75.0	69	4	US-08-856-841-18	Sequence 18, Appl
42	27	75.0	69	4	US-08-856-841-18	Sequence 18, Appl
43	27	75.0	74	2	US-08-856-841-18	Sequence 18, Appl
44	27	75.0	74	2	US-08-856-841-18	Sequence 18, Appl
45	27	75.0	89	3	US-08-856-841-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-09-134-001C-3833
Sequence 3833, Application US/09134001
Patent No. 6380370
GENERAL INFORMATION:
AFFILIANT: 1000 Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR STAPHYLOCOCCUS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT FILING DATE: 1998-08-13
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3833
LENGTH: 204
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3833

Query Match 88.9% Score 127 DB 4 Length 204
Best Local Similarity 71.4% Pred. No. 127
Matches 51 Conservative 21 Mismatches 0 Indels 0 Gaps 0

OR 1 KEYPTS 7
DB 195 VKEYPTA 201

RESULT 2
US-09-134-001C-3015
Sequence 3015, Application US/09134001
Patent No. 6380370
GENERAL INFORMATION:
AFFILIANT: 1000 Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR STAPHYLOCOCCUS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT FILING DATE: 1998-08-13
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5574
SEQ ID NO 3015
LENGTH: 703
TYPE: PRT

ORGANISM: Staphylococcus epidermidis
09-134-001C-3015

Query Match 80.6% Score 29; DB 4; Length 703;
est Local Similarity 83.3%; Pred. No. 2,1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

2 KEYS 7

1:1111
195 KDFTS 200

ULT 4
08-856-841-9
Sequence 9, Application US/08856841
Patent No. 6214307

GENERAL INFORMATION:

APPLICANT: ERWIN SOUTSCHER
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS R19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY

COUNTRY: USA
ZIP: 10016

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER US/08/856,841
FILING DATE:

CLASSIFICATION:

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US/08/214,659
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/US91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LYP-922-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635

TELEX: NONE

INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:

LENGTH: 210

TYPE: AMINO ACID

TOPOLOGY: LINEAR

MOLECULE TYPE: PEPTIDE

DESCRIPTION: N/A

ANTI-SENSE: N/A

FRAGMENT TYPE: INTERNAL

ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE

IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE

POSITION IN GENOME: N/A

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD: amino acid analysis and

IDENTIFICATION METHOD: mass spectrometry

OTHER INFORMATION:

PUBLICATION INFORMATION:

AUTHORS: COSSART, Y. E.

AUTHORS: FIELD, A. M.

AUTHORS: GANT, B.

AUTHORS: WIDOWS, D.

TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA

JOURNAL: LANCET

VOLUME: 1

ISSUE: 72 - 73

PAGES: 1975

DATE: 1975

DOCUMENT NUMBER:

FILING DATE:

PUBLICATION DATE:

RELEVANT RESIDUES IN SEQ ID NO: 9:

PUBLICATION INFORMATION:

AUTHORS: MANIATIS, T.

AUTHORS: FRITSCH, E. F.

AUTHORS: SAMBROOK, J.

TITLE: MOLECULAR CLONING

JOURNAL: COLD SPRING HARBOR, NY

VOLUME:

ISSUE:

PAGES:

DATE: 1982

DOCUMENT NUMBER:

FILING DATE:

PUBLICATION DATE:

RELEVANT RESIDUES IN SEQ ID NO: 9:

PUBLICATION INFORMATION:

AUTHORS: SMITH, D. B.

AUTHORS: THOMPSON, K. S.

TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES

TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH

TITLE: GLUTATHIONE S- TRANSFERASE

JOURNAL: GENE

VOLUME:

ISSUE: 67

PAGES: 31 - 40

DATE: 1988

DOCUMENT NUMBER:

FILING DATE:

PUBLICATION DATE:

RELEVANT RESIDUES IN SEQ ID NO: 9:

US-08-856-841-9

Query Match 77.8% Score 29; DB 4; Length 210;

est Local Similarity 66.7%; Pred. No. 92;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

DB 194 KDFT 199

RESULT 4

Sequence 3, Application US/0881583

Patent No. 6218142

GENERAL INFORMATION:

APPLICANT: Wasseneget, Michael

APPLICANT: Riedel, Leonhard

APPLICANT: Schiebel, Winfried

APPLICANT: Saenger, Heinz

TITLE OF INVENTION: NUCLEIC ACID MIMETICS ENCODING

TITLE OF INVENTION: POLYPEPTIDES HAVING THE ENZYMAIC ACTIVITY OF AN

TITLE OF INVENTION: RNA-DIRECTED RNA POLYMERASE (RDRP)

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: FISH & NEAVE

STREET: 1251 Avenue of the Americas

CITY: New York

STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-09/647-749a
FILING DATE: 05-MAR-1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Haley, James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: MP3-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 218 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-583-3

Query Match: 77.8% Score 28; DR 4; Length 218;
Best Local Similarity: 57.1% Pred. No. 96;
Matches: 4; Conservative: 3; Mismatches: 0; Indels: 0; Gaps: 0

1 KEYS: 7
144 VEIPTN 150

RESULT 5
US-08-856-841-15
Sequence: 15; Application US/08856841
Patent No. 6274407

GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHER
APPLICANT: MANFRED MOTT
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OF POLYPEPTIDES FROM THE PAPOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-09/647-749a
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US-08/214,656
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/CE91/00102
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:

NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9227-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 227
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: N/A
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (EPSTEIN-BARR VIRUS INFECTION)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSAFT, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PAPOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 15:
PUBLICATION INFORMATION:
AUTHORS: MANIATIS, T.
AUTHORS: PRITSCH, F.F.
AUTHORS: SAMBROOK, J.
TITLE: MOLECULAR CLONING
JOURNAL: COLD SPRING HARBOR, NY
VOLUME:
ISSUE:
PAGES:
DATE: 1982
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 17:
PUBLICATION INFORMATION:
AUTHORS: SMITH, D.B.
AUTHORS: JOHNSON, K.S.
TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
JOURNAL: GUTATHIONE S. TRANSFERASE
VOLUME:
ISSUE:
PAGES: 67
DATE: 31 - 40
FILING DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 17:
US-08-856-841-15

Query Match 77.8% Score 28 DB 4 Length 257
 Best Local Similarity 66.7% Pred. NO. 1e+02
 Matches 4 Conservative 2 Mismatches 0 Indels 0 Gaps 0

1 KEYFT 6
 11111
 193 VKDYFT 198

SULT 6
 -08-856-841-12
 Sequence 12, Application US/08856841
 Patent No. 6274307
 GENERAL INFORMATION:

APPLICANT: MANFRED MOTZ
 TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
 TITLE OF INVENTION: CP POLYPEPTIDES FROM THE PARVOVIRUS S19
 NUMBER OF SEQUENCES: 28
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BROOKS BAIRD HAFNER & DELAUNTY
 STREET: 96 PARK AVENUE
 CITY: NEW YORK
 STATE: NY
 COUNTRY: USA
 ZIP: 10016

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" FLOPPY DISC
 COMPUTER: AT&T - IBM COMPATIBLE
 OPERATING SYSTEM: MS-DOS VERSION 6.2
 SOFTWARE: ASCII

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US-08-856-841
 FILING DATE:
 CLASSIFICATION:

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US-08-214,658
 FILING DATE: 16-MARCH-1994
 APPLICATION NUMBER: US 07/917,096
 FILING DATE: 4-AUGUST-1992
 APPLICATION NUMBER: PCT/DE91/00106
 FILING DATE: 8-FEBRUARY-1991
 APPLICATION NUMBER: DE40038262
 FILING DATE: 8-FEBRUARY-1990

ATTORNEY/AGENT INFORMATION:
 NAME: ROBINSON, WILLIAM R.
 REGISTRATION NUMBER: 27,224
 REFERENCE/COCKET NUMBER: LKF-0000-A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 697-3355
 TELEFAX: (212) 557-5635

TELEX: NONE
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:

LENGTH: 250
 TYPE: AMINO ACID
 TOPOLOGY: LINEAR
 MOLECULE TYPE:
 DESCRIPTION: PEPTIDE
 ANTI-SENSE: N/A
 HYPOTHETICAL: N/A
 PRAGMENT TYPE: INTERNAL
 ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
 ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
 IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
 POSITION IN GENOME: N/A
 FEATURE:

NAME/KEY: N/A
 LOCATION: N/A
 IDENTIFICATION METHOD: amino acid analysis and
 IDENTIFICATION METHOD: mass spectrometry
 OTHER INFORMATION:
 PUBLICATION INFORMATION:

AUTHORS: COSGART, Y.E.
 AUTHORS: FIELD, A.M.
 AUTHORS: CANT, B.
 AUTHORS: WIDOWS, D.
 TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
 JOURNAL: LANCET
 VOLUME: 1
 ISSUE:
 PAGES: 72 - 73
 DATE: 1975

DOCUMENT NUMBER:
 FILING DATE:
 PUBLICATION DATE:
 RELEVANT RESIDUES IN SEQ ID NO: 12:
 PUBLICATION INFORMATION:

AUTHORS: MANIATIS, T.
 AUTHORS: FRITSCH, E.F.
 AUTHORS: SAMBROOK, J.
 TITLE: MOLECULAR CLONING
 JOURNAL: COLD SPRING HARBOR, NY
 VOLUME:
 ISSUE:
 PAGES:
 DATE: 1982

DOCUMENT NUMBER:
 FILING DATE:
 PUBLICATION DATE:
 RELEVANT RESIDUES IN SEQ ID NO: 12:
 PUBLICATION INFORMATION:

AUTHORS: SMITH, D.B.
 AUTHORS: JOHNSON, K.S.
 TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES WITH
 TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
 TITLE: GLUTATHIONE S- TRANSFERASE
 JOURNAL: GENE
 VOLUME:
 ISSUE: 67
 PAGES: 31 - 40
 DATE: 1988

DOCUMENT NUMBER:
 FILING DATE:
 PUBLICATION DATE:
 RELEVANT RESIDUES IN SEQ ID NO: 12:
 US-08-856-841-12

Query Match 77.8% Score 28 DB 4 Length 250
 Best Local Similarity 66.7% Pred. NO. 1e+02
 Matches 4 Conservative 2 Mismatches 0 Indels 0 Gaps 0

1 KEYFT 6
 11111
 DB 200 VKDYFT 205

RESULT 7
 US-08-935-263-16
 Sequence 16, Application US/08935263A
 Patent No. 6117669

GENERAL INFORMATION:
 APPLICANT: Furutachi, Yasuhiro
 APPLICANT: Hoshino, Tatsuo
 APPLICANT: Kimura, Hitoshi
 APPLICANT: Kiyasu, Tatsuya
 APPLICANT: Nagashima, Yoshie
 TITLE OF INVENTION: BIOTIN BIOSYNTHETIC GENES
 FILE REFERENCE: Biotin Genes
 CURRENT APPLICATION NUMBER: US/08/935,263A
 CURRENT FILING DATE: 1997-09-22
 EARLIER APPLICATION NUMBER: EP 96115540.5
 EARLIER FILING DATE: 1996-09-27
 NUMBER OF SEQ ID NOS: 23
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 16

LENGTH: 276
 TYPE: PRT
 ORGANISM: Kurthia sp.
 IS-08-935-263-16

Query Match 77.8% Score 28; DB 3; Length 276
 Best Local Similarity 57.1% Pred. No. 1,2e+02;
 Matches 4; Conservative 3; Mismatches 0; Indels 0;

1 IKEYFS 7
 214 VREFFS 220

RESULT 9
 IS-09-594-185-16
 Sequence 16, Application US/08/054195
 Patent No. 6165388
 GENERAL INFORMATION:

APPLICANT: Furutachi, Yasuhiro
 APPLICANT: Hoshino, Tatsuo
 APPLICANT: Kimura, Hitoshi
 APPLICANT: Kiyasu, Tatsuya
 APPLICANT: Nagahashi, Yoshie
 TITLE OF INVENTION: BIOTIN BIOSYNTHETIC GENES
 FILE REFERENCE: Biotin Genes
 CURRENT APPLICATION NUMBER: US/09/764,185
 PRIOR FILING DATE: 2000-06-14
 PRIOR APPLICATION NUMBER: 08/935,263
 PRIOR FILING DATE: 1997-09-22
 PRIOR APPLICATION NUMBER: EP 96115540.5
 PRIOR FILING DATE: 1996-09-27
 NUMBER OF SEQ ID NOS: 23
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 16

LENGTH: 276
 TYPE: PRT
 ORGANISM: Kurthia sp.
 IS-09-594-185-16

Query Match 77.8% Score 28; DB 4; Length 276;
 Best Local Similarity 57.1% Pred. No. 1,2e+02;
 Matches 4; Conservative 3; Mismatches 0; Indels 0;

1 IKEYFS 7
 214 VREFFS 220

RESULT 9
 S-08-605-284B-18
 Sequence 18, Application US/08/05284B
 Patent No. 6060271
 GENERAL INFORMATION:

APPLICANT: WALEWSKI, JOSE F
 APPLICANT: RECIO-PINTO, ESPERANZA
 TITLE OF INVENTION: VOLTAGE GATED SODIUM CHANNELS FROM
 TITLE OF INVENTION: HUMAN PERIPHERAL NERVE
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
 STREET: CLINTON SQUARE, P.O. BOX 1051
 CITY: ROCHESTER
 STATE: NEW YORK
 COUNTRY: USA

ZIP: 14603
 COMPUTER READABLE FORM
 MEDIUM TYPE: FLOPPY disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/605,284B

FILING DATE: 09-FEB-1996
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: BRAMAN, SUSAN J.
 REGISTRATION NUMBER: 34,103
 REFERENCE/EXCERPT NUMBER: 1996-08-01 (OFF D-1705)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 716-263-1636
 TELEFAX: 716-263-1600
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 311 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-605-284B-18

Query Match 77.8% Score 28; DB 3; Length 311;
 Best Local Similarity 71.4% Pred. No. 1,4e+02;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

1 IKEYFS 7
 147 IKEYFS 153

RESULT 10
 US-08-856-841-17
 Sequence 17, Application US/08/856841
 Patent No. 6274307
 GENERAL INFORMATION:

APPLICANT: ERWIN SOUTSCHER
 APPLICANT: MANFRED MOLT
 TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
 TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
 NUMBER OF SEQUENCES: 28
 CORRESPONDENCE ADDRESS:

ADDRESSEE: BROOKS HAUPT HAFNER & DELAHOUDY
 STREET: 99 PARK AVENUE
 CITY: NEW YORK
 STATE: NY
 COUNTRY: USA
 ZIP: 10016

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" FLOPPY DISC
 COMPUTER: AT&T - IBM COMPATIBLE
 OPERATING SYSTEM: MS-DOS Version 6.2
 SOFTWARE: ASCII

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/856,841
 FILING DATE:

CLASSIFICATION:
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/214,654
 FILING DATE: 16-MARCH-1994
 APPLICATION NUMBER: US 07/917,054
 FILING DATE: 4-AUGUST-1992
 APPLICATION NUMBER: PT/96/01,001
 FILING DATE: 8-FEBRUARY-1991
 APPLICATION NUMBER: DE40038262
 FILING DATE: 8-FEBRUARY-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: ROBINSON, WILLIAM R.
 REGISTRATION NUMBER: 27,224
 REFERENCE/DOCKET NUMBER: LRP 922,00A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 697-3355
 TELEFAX: (212) 557-5635
 TELEX: NONE

INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 387

TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: N/A
HYPOHEMICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (EPITHEMIA INFECTION)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDOWS, D.
AUTHORS: FRITSCHE, E.F.
AUTHORS: SAMROCK, J.
TITLE: MOLECULAR CLONING
JOURNAL: COLD SPRING HARBOR, NY
VOLUME:
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 17:
PUBLICATION INFORMATION:
AUTHORS: MANIATIS, T.
AUTHORS: FRITSCHE, E.F.
AUTHORS: SAMROCK, J.
TITLE: MOLECULAR CLONING
JOURNAL: COLD SPRING HARBOR, NY
VOLUME:
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 15:
PUBLICATION INFORMATION:
AUTHORS: SMITH, D.B.
AUTHORS: JOHNSON, K.S.
TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
TITLE: GLUTATHIONE S. TRANSFERASE
JOURNAL: GENE
VOLUME:
ISSUE: 67
PAGES: 31 - 40
DATE: 1988
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 17:
-08-856-841-17
Query Match 77.88; Score 28, DB 4; Length 387;
Best Local Similarity 66.78; Freq. No. 1.8e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
1 KEYFT 6
1:111
201 VKDYFT 206
SULT 11

US-08-856-841-19
Sequence 19, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHER
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESS: BROOKS HAIRY HAFNER & DELAUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/08/214,654
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,094
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00100
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-922,77A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 486
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: N/A
HYPOHEMICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (EPITHEMIA INFECTION)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDOWS, D.
AUTHORS: FRITSCHE, E.F.
AUTHORS: SAMROCK, J.
TITLE: MOLECULAR CLONING
JOURNAL: COLD SPRING HARBOR, NY
VOLUME:
ISSUE:
PAGES: 72 - 73
DATE: 1975

DOCUMENT NUMBER:
 FILING DATE:
 PUBLICATION DATE:
 RELEVANT RESIDUES IN SEQ ID NO: 19:
 S-08-856-841-19

Query Match 77.88; Score 28; DB 4; Length 486;
 Best Local Similarity 66.78; Pred. N. 2.3e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Y 1 KEYFT 6
 D 193 VKDYFT 198

RESULT 12
 S-08-856-841-16
 Sequence 16 Application US/08866941
 Patent No. 6274307
 GENERAL INFORMATION:

APPLICANT: ERWIN SOUTSCHER
 APPLICANT: MANFRED MOTZ
 TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
 TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVAVIRUS P13
 NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:
 ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
 STREET: 99 PARK AVENUE
 CITY: NEW YORK
 STATE: NY
 COUNTRY: USA
 ZIP: 10016

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" FLOPPY DISC
 COMPUTER: AT&T - IBM COMPATIBLE
 OPERATING SYSTEM: MS-DOS Version 6.2
 SOFTWARE: ASCII

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US-08-856-841
 FILING DATE:

CLASSIFICATION:
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: US-08-8214-556
 FILING DATE: 16-MARCH-1994
 APPLICATION NUMBER: US-07-917-036
 FILING DATE: 4-AUGUST-1992
 APPLICATION NUMBER: PCT/DE91/000106
 FILING DATE: 8-FEBRUARY-1991
 APPLICATION NUMBER: DE00038262
 FILING DATE: 8-FEBRUARY-1990

ATTORNEY/AGENT INFORMATION:
 NAME: ROBINSON, WILLIAM R.
 REGISTRATION NUMBER: 27,224
 REFERENCE NUMBER: 17-00000A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 697-3355
 TELEFAX: (212) 557-5635
 TELEX: NONE

INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:

LENGTH: 500
 TYPE: AMINO ACID
 TOPOLOGY: LINEAR
 MOLECULE TYPE:
 DESCRIPTION: PEPTIDE
 HYPOTHETICAL: N/A
 ANTI-SENSE: N/A
 FRAGMENT TYPE: INTERNAL
 ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
 IMMEDIATE SOURCE: INFECTION (EPITHEMA INFECTIONUM)
 IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
 POSITION IN GENE: N/A
 FEATURE:

NAME/KEY: N/A
 LOCATION: N/A
 IDENTIFICATION METHOD: amino acid analysis and
 IDENTIFICATION METHOD: mass spectrometry
 OTHER INFORMATION:

PUBLICATION INFORMATION:
 AUTHORS: COSSART, Y.E.
 AUTHORS: FIELD, A.M.
 AUTHORS: CANT, B.
 AUTHORS: WILKINS, D.
 TITLE: PARVAVIRUS-LIKE PARTICLES IN HUMAN SERA
 JOURNAL: LANCET
 VOLUME: 1
 ISSUE:
 PAGES: 72 - 73
 DATE: 1975

DOCUMENT NUMBER:
 FILING DATE:
 PUBLICATION DATE:
 RELEVANT RESIDUES IN SEQ ID NO: 16:
 PUBLICATION INFORMATION:

AUTHORS: MANIATIS, T.
 AUTHORS: FRITSCH, E.F.
 AUTHORS: SAMPOUR, J.
 TITLE: MOLECULAR CLONING
 JOURNAL: COLD SPRING HARBOR, NY
 VOLUME:
 ISSUE:
 PAGES:
 DATE: 1982

DOCUMENT NUMBER:
 FILING DATE:
 PUBLICATION DATE:
 RELEVANT RESIDUES IN SEQ ID NO: 16:
 PUBLICATION INFORMATION:

AUTHORS: SMITH, D.B.
 AUTHORS: JOHNSON, K.S.
 TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
 TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
 TITLE: GLUTATHIONE S. TRANSFERASE
 JOURNAL: GENE
 VOLUME:
 ISSUE: 67
 PAGES: 31 - 40
 DATE: 1988

DOCUMENT NUMBER:
 FILING DATE:
 PUBLICATION DATE:
 RELEVANT RESIDUES IN SEQ ID NO: 16:
 US-08-856-841-16

Query Match 77.88; Score 28; DB 4; Length 500;
 Best Local Similarity 66.78; Pred. N. 2.4e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Y 1 KEYFT 6
 D 201 VKDYFT 206

RESULT 13
 US-08-856-841-18
 Sequence 18 Application US/08866841
 Patent No. 6274307

GENERAL INFORMATION:
 APPLICANT: ERWIN SOUTSCHER
 APPLICANT: MANFRED MOTZ
 TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
 TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVAVIRUS P13
 NUMBER OF SEQUENCES: 28
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
 STREET: 99 PARK AVENUE

CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/04/956,841
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/04/214,678
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/EP/93/001,006
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LAR-0222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ. ID NO.: 1a:
SEQUENCE CHARACTERISTICS:
LENGTH: 501
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
IMMEDIATE SOURCE: INFECTION (ERYTHEMA INFEKTIOSUM)
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y. E.
AUTHORS: FLETH, A. M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO.: 18:
-08-856-841-18
Query Match 77.8% Score 28; DB 4; Length 501;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
1 IKEYFT 6
|||||

DB 202 VKDYFT 207

RESULT 14
US-09-413-452-2
Sequence 2, Application US/09413452
Patent No. 6083540

GENERAL INFORMATION:
APPLICANT: Christensen, T.
APPLICANT: Thorsoe, H.
APPLICANT: Kreiberg, J.
APPLICANT: Buchholt, H.
TITLE OF INVENTION: PROCESS FOR STABILIZING PROTEINS IN AN
ACIDIC ENVIRONMENT WITH A HIGH-ESTER PECTIN
FILE REFERENCE: DYN014 001APC
CURRENT APPLICATION NUMBER: US/09/413,452
FILING DATE: 1999-10-06
EARLIER APPLICATION NUMBER: 06/983364
FILING DATE: 1998-05-18
EARLIER APPLICATION NUMBER: PCT/EP96/0051
FILING DATE: 1996-07-12
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FASTSEQ for Windows Version 1.0
SEQ ID NO 2
LENGTH: 584
TYPE: PPT
ORGANISM: Citrus navelina var. class 1
US-09-413-452-2

Query Match 77.8% Score 28; DB 3; Length 584;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKEYFT 6
|||||
DB 4 IKEYFT 9

RESULT 15
US-09-413-068-2
Sequence 2, Application US/09413068
Patent No. 6268195

GENERAL INFORMATION:
APPLICANT: Christensen, T.
APPLICANT: Thorsoe, H.
APPLICANT: Kreiberg, J.
APPLICANT: Buchholt, H.
TITLE OF INVENTION: PROCESS FOR STABILIZING PROTEINS IN AN
ACIDIC ENVIRONMENT WITH A HIGH-ESTER PECTIN
FILE REFERENCE: DYN014 001APC
CURRENT APPLICATION NUMBER: US/09/413,068
FILING DATE: 1999-06-06
EARLIER APPLICATION NUMBER: 08/983364
FILING DATE: 1998-05-18
EARLIER APPLICATION NUMBER: PCT/EP96/0051
FILING DATE: 1996-07-12
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FASTSEQ for Windows Version 1.0
SEQ ID NO 2
LENGTH: 584
TYPE: PPT
ORGANISM: Citrus navelina var. class 1
US-09-413-068-2

Query Match 77.8% Score 27; DB 4; Length 584;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKEYFT 6
|||||
DB 4 IKEYFT 9

Tue Apr 8 16:02:02 2003

us-09-647-749a-2.ra1

Page 9

Search completed: April 8, 2003, 12:05:53
Job time : 23.5833 Secs

|||||
367 IKEYFT 372

SUIT 2
CV-CHICK
PGCV-CHICK STANDARD: PRT: 3562 AA.

090953: 090945: 01-NOV-1997 (Rel. 35, Created)

01-NOV-1997 (Rel. 35, Last sequence update)

15-JUN-2002 (Rel. 41, Last annotation update)

Version core protein precursor (large fibroblast proteoglycan)

(Chondroitin sulfate proteoglycan core protein 2) (PG-M).

CP562.
Gallus gallus (Chicken).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
Gallus.

NCBI_TaxID=9031;

(1)

SEQUENCE FROM N.A.
STRAIN-White Leghorn; TISSUE=Limb bud;

MEDLINE=93300846; PubMed=8314802;

Shinomura T., Nishida Y., Ito K., Kimura K.;

cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan
expressed during chondrogenesis in chick limb buds. Alternative
spliced multiforms of PG-M and their relationships to versican."

J. Biol. Chem. 268:14461-14469(1993).

-1- FUNCTION: May play a role in intercellular signaling and in
connecting cells with the extracellular matrix. May take part in
the regulation of cell motility, growth and differentiation. Binds
hyaluronic acid.

-1- SUBCELLULAR LOCATION: Secreted; extracellular matrix.

-1- ALTERNATIVE PRODUCTS: At least 2 isoforms: V1 (shown here) and
V2: are produced by alternative splicing.

-1- TISSUE SPECIFICITY: Prechondrogenic condensation area of
developing limb buds.

-1- DEVELOPMENTAL STAGE: Disappears after the cartilage development
(by similarity).

-1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.

-1- SIMILARITY: CONTAINS 2 LINK DOMAINS.

-1- SIMILARITY: CONTAINS 4 EGF-LIKE DOMAINS.

-1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.

-1- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.

This SWISS-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb.ac.uk/embout>
or send an email to license@isb.ac.uk).

EMBL: X60236; CAA42787.1;

EMBL: D13542; BA02742.1;

HSSP: P00743; IELM.

InterPro: IPR000152; ASX-Hydroxyl.

InterPro: IPR000561; EGF-like.

InterPro: IPR000742; EGF-2.

InterPro: IPR001861; EGF-3.

InterPro: IPR003599; IG-1.

InterPro: IPR003906; IG-2.

InterPro: IPR001304; LECTIN-C.

InterPro: IPR000538; LINK.

InterPro: IPR000435; SUSHI_SCR-CDP.

Pfam: PF000047; IGF_1.

Pfam: PF00059; IGF_1.

Pfam: PF00084; SUSHI_1.

Pfam: PF00193; XLINK_2.

ProDom: PD000918; LINK_2.

SMART: SM00032; CCP_1.

SMART: SM00034; CECT_1.

DR SMART: SM00179; EGF_CA_1.

DR SMART: SM00001; EGF-Like_1.

DR SMART: SM00409; IG_1.

DR SMART: SM00445; LINK_2.

DR PROSITE: PS00010; ASX-Hydroxyl_1.

DR PROSITE: PS00022; EGF-1_2.

DR PROSITE: PS01186; EGF-2_1.

DR PROSITE: PS01187; EGF-CA_1.

DR PROSITE: PS01241; LINK_2.

DR PROSITE: PS00615; C-TYPE-LECTIN-1.

DR PROSITE: PS00041; C-TYPE-LECTIN-2.

DR Glycoprotein: Proteoglycan; lectin; Extracellular matrix; SUSHI;

Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;

KW Hyaluronic acid; Alternative splicing.

KW SIGNAL

FT CHAIN 27 3562

FT DOMAIN 37 136

FT DOMAIN 166 243

FT DOMAIN 264 345

FT DOMAIN 3254 3290

FT DOMAIN 3292 3328

FT DOMAIN 3341 3455

FT DOMAIN 3460 3518

FT DISULFID 44 129

FT DISULFID 171 242

FT DISULFID 195 216

FT DISULFID 269 344

FT DISULFID 293 314

FT DISULFID 3258 3269

FT DISULFID 3263 3278

FT DISULFID 3280 3289

FT DISULFID 3296 3307

FT DISULFID 3301 3316

FT DISULFID 3318 3327

FT DISULFID 3334 3345

FT DISULFID 3362 3454

FT DISULFID 3430 3446

FT DISULFID 3461 3504

FT DISULFID 3490 3517

FT CARBOHYD 152 162

FT CARBOHYD 235 235

FT CARBOHYD 329 329

FT CARBOHYD 529 529

FT CARBOHYD 709 709

FT CARBOHYD 948 948

FT CARBOHYD 1406 1406

FT CARBOHYD 1479 1479

FT CARBOHYD 1523 1523

FT CARBOHYD 1530 1530

FT CARBOHYD 1545 1545

FT CARBOHYD 1571 1571

FT CARBOHYD 1968 1968

FT CARBOHYD 2088 2088

FT CARBOHYD 2089 2089

FT CARBOHYD 2507 2507

FT CARBOHYD 2642 2642

FT CARBOHYD 2679 2679

FT CARBOHYD 2748 2748

FT CARBOHYD 2762 2762

FT CARBOHYD 3069 3069

FT CARBOHYD 3194 3194

FT CARBOHYD 3232 3232

FT CARBOHYD 3545 3545

FT VARPELIC 485 1411

SV SEQUENCE 3562 AA; 388078 MW; 945565E89216222 CP064;

Query Match

Best Local Similarity 100.0%;

Matches 6; Conservative 0; Mismatches 0;

Of 2 KEYFTS 7

DB 1355 KEYFTS 1360

Score 32; DB 1; Length 3562;

Pred. No. 1.7e+02;

Indels 0; Gaps 0;

```

RESULT 3
TC1_DICD1
D RTCL_DICD1 STANDARD: PRT: 433 AA.
C 015746:
T 15-DEC-1998 (Rel. 37, Last sequence update)
T 16-OCT-2001 (Rel. 40, Last annotation update)
T 16-OCT-2001 (Rel. 40, Last annotation update)
E Probable RNA 3'-terminal phosphate cyclase (EC 6.5.1.4) (PNA-3'-phosphate cyclase) (RNA cyclase).
N R2037.
S Dictyostelium discoideum (Slime mold)
C Eukaryota: Microsporidia: Eukaryotina: Dictyostelium
X NCBI_TaxID=44689.
P SEQUENCE FROM N.A.
(1)
D Loomis W.F., Tranter N.
C STRAIN=AX3
C Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases
C -1- FUNCTION: CATALYZES THE CONVERSION OF 3'-PHOSPHATE TO A 3',3'-CYCLIC PHOSPHODIESTER AT THE END OF RNA. THE MECHANISM OF ACTION OF THE ENZYME OCCURS IN 3 STEPS: (A) ADENYLATION OF THE ENZYME BY ATP; (B) THE ENZYME ACTS ON RNA-N3'P TO PRODUCE RNA-N3'PP5'A; (C) A NON CATALYTIC NUCLEOPHILIC ATTACK BY THE ADJACENT 2' HYDROXYL ON THE PHOSPHORUS IN THE FIRST STEP LINKAGE TO PRODUCE THE CYCLIC END-PRODUCT. THE BIOLOGICAL ROLE OF THIS ENZYME IS UNKNOWN BUT IT IS LIKELY TO FUNCTION IN SOME ASPECTS OF CELLULAR RNA PROCESSING (BY SIMILARITY).
C -1- CATALYTIC ACTIVITY: ATP + RNA 3'-terminal-phosphate -> AMP + diphosphate + RNA terminal-2',3'-cyclic-phosphate.
C -1- SUBCELLULAR LOCATION: Nuclear; nucleoplasm (By similarity).
C -1- SIMILARITY: BELONGS TO THE RNA 3'-TERMINAL CYCLASE FAMILY.
C -1- SIMILARITY: 1.
C -1- CAUTION: THE C-TERMINAL SECTION OF THIS PROTEIN COULD BE INCOMPLETE AND THE REAL PROTEIN IS PROBABLY SHORTER.
C
C THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed, usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch or send an email to license@sib-sib.ch).
C
R EMBL: AF020279; AAB70847.1; -.
R DictyDB: DP02222; -.
R InterPro: IPR000228; PNAS_term_cycl.
R Pfam: PF01137; RTC1.1.
R ProSite: PS01289; RTC1.1.
W Ligase: Nuclear protein.
T ACT_SITE 354 354 BY SIMILARITY
Q SEQUENCE 433 AA: 47110 MW: 71496044458699 GRC64:
Query Match 86.1% Score 31; DB 1; Length 433;
Best Local Similarity 85.7% Pred. No. 30;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Z 1 KEYSPTS 7
b 229 LKEYPTS 234

```

```

OC Eukaryota, Fungi, Ascomycota, Saccharomycotina, Saccharomycetes;
OC Saccharomycetales, Saccharomycetaceae, Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
P SEQUENCE FROM N.A.
R MEDLINE=9526475; PubMed=7744966;
R Gersch L.C., Corkendorff T.C., Collier N.;
T "A conditional allele of the novel repeat-containing yeast nucleoporin NUP159 causes both rapid cessation of mRNA export and reversible clustering of nuclear pore complexes."
J. Cell Biol. 129:939-956(1995).
N
(2)
P SEQUENCE FROM N.A.
P STRAIN=5289; AP072;
R Baird B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
R Chatterjee C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
R Gentes S., Hamlyn N., Horsnell T.S., Hunt S., Jagals K., Jones M.,
R Louis E., Ly G., Moulé S., Moulé T., Odell C., Pearson D.,
R Patandream M.A., Piles L., Powley N., Skelton J., Smith V.,
R Walsh S.V., Whitehead S.;
T Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
C -1- FUNCTION: MAY PLAY A DIRECT ROLE IN NUCLEOTIDYLTRANSFERASE EXPORT OF RNA AND ALSO IN PROTEIN IMPORT.
C -1- SUBUNIT: INTERACTS WITH DBP5.
C -1- SUBCELLULAR LOCATION: Nuclear pore complex.
C -1- DOMAIN: CONTAINS A CENTRAL REGION WITH REPEATS THAT ARE X-X-F-G, PRIMARILY S-A-F-G AND P-S-F-G. SOME OF THESE REPEATS ARE EMBEDED WITH IN A NEARLY PERFECT 26 AMINO ACID TANDEM REPEATS.
C
C THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed, usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
C
R EMBL: L40634; AAC41652.1; -.
R EMBL: Z8125; CA86265.1; -.
R PIR: S48457; S48457.
R S300: S0001377; NUP159.
R InterPro: IPR004325; Nucleoporin_FG
R Pfam: PF01794; Nucleoporin_FG
R Kinase: K117; 13.
R Nuclear protein; Transports; Repeat
R DOMAIN 518 617
R REPEAT 518 543 1.
R REPEAT 544 569 2.
R REPEAT 570 595 3.
R REPEAT 596 617 4.
R SEQUENCE 1460 AA: 158907 MW: 441945237858205 GRC64:
Query Match 86.1% Score 41; DB 1; Length 1460;
Best Local Similarity 71.4% Pred. No. 11602;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Z 1 KEYSPTS 7
Db 1190 LKEYPTS 1196

```

```

RESULT 5
PTKA_ECOLI
ID PTKA_ECOLI STANDARD: PRT: 150 AA.
AC P37187; P76413;
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE PTS system, galactitol-specific IIA component (Eukaryot-
DE per-mease IIA component) (Phosphotransferase enzyme II A component)
DE (EC 2.7.1.69).
DE GATA OR B2094 OR Z3257 OR FCS2897.
OS Escherichia coli, and
OS Escherichia coli O157:H7.

```


Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae.

Escherichia

NCBI_TaxID=562, 83334.

[1]

SEQUENCE FROM N.A.

STRAIN-EC3132

MEDLINE=95290477; PubMed=7772643;

Nobelmann B., Lenzeler J.W.;

*Sequence of the gat operon for galactitol utilization from a

wild-type strain EC3132 of *Escherichia coli*.*

Biochim. Biophys. Acta 1262:69-72(1995).

[2]

SEQUENCE FROM N.A.

STRAIN-K12 / MG1655

MEDLINE=97426617; PubMed=9276503;

Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

Mau B., Shao Y.;

*The complete genome sequence of *Escherichia coli* K-12.*

Science 277:1232-1244(1997).

[3]

SEQUENCE FROM N.A.

STRAIN-O157:H7 / ATCC 35762

MEDLINE=21074935; PubMed=1126551;

Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,

Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,

Postal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,

Goebcke E.J., Davis N.W., Lim A., Dolananta E.T., Potamouis K.,

Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,

Weich R.A., Blattner F.R.;

*Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.*

Nature 409:529-533(2001).

[4]

SEQUENCE FROM N.A.

STRAIN-O157:H7 / RMD 0509952

MEDLINE=21156231; PubMed=11258796;

Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokogawa K.,

Han C.-G., Ohtsuo E., Nakayama K., Murata T., Tanaka M., Toke T.,

Iida T., Takami H., Honda T., Sasakawa G., Ogasawara N., Yasunaga T.,

Kudoh S., Shiba T., Hattori M., Shiwagawa H.;

*Complete genome sequence of enterohaemorrhagic *Escherichia coli*

O157:H7 and genomic comparison with a laboratory strain K-12.*

DNA Res. 8:11-22(2001).

-1- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHONOPYRUVATE-DEPENDENT

SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS). A MAJOR CARBOHYDRATE ACTIVE

TRANSPORT SYSTEM. THE ITC DOMAIN CONTAINS THE SUGAR BINDING SITE

AND THE TRANSMEMBRANE CHANNEL. THE IIA DOMAIN CONTAINS THE PRIMARY

PHOSPHORYLATION SITE (THE DOMAIN IS PHOSPHO-HRP); IIA TRANSFERS ITS

PHOSPHORYL GROUP TO THE IIB DOMAIN WHICH FINALLY TRANSFERS IT TO

THE SUGAR.

-1- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar -> protein

histidine + sugar phosphate

-1- SUBCELLULAR LOCATION: Cytoplasmic.

-1- SIMILARITY: CONTAINS 1 PTS IIA DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/submitter/> or send an email to license@sib-sib.ch).

EMBL: X79837; CAA56228.1;

EMBL: AE000298; AAC75155.1;

EMBL: AE005435; AAG57511.1;

EMBL: AP002560; RA046320.1;

Ecogene: EG12414; gata.

InterPro: IPR002178; PTS_EIIA-2.

Protein: PF00359; PTS_EIIA_2; 1.

Protein: P0001689; PTS_EIIA_2; 1.

Phosphotransferase system: Sugar transport: Transferase:

KM Phosphorylation; Galactitol metabolism; Complete proteome.

FT CONFLICT 124 124 D -> E (IN REF. 1).

SC SEQUENCE 150 AA; 16907 MW; 1A760076B5FB61D CRC64.

Query Match

Best Local Similarity 83.3%; Pred. No. 15;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKEYFT 6

DB 1 IKEYFT 142

DB 1 IKEYFT 142

RESULT 6

PYBL_PEA

ID PYBL_PEA

AC 043086;

DT 15-DEC-1998 (rel. 37, Created)

DT 15-DEC-1998 (rel. 37, Last sequence update)

DT 15-JUN-2002 (rel. 41, Last annotation update)

DE Aspartate carbamoyltransferase 1, chloroplast precursor (EC 2.3.1.2)

DE (Aspartate transcarbamylase 1) (AtCase 1).

GN PYBL.

OS Pisum sativum (Garden pea).

OC Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta;

OC Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots, Rosidae;

OC eucosids I, Fabales, Fabaceae, Papilionoideae, Viciae; Pisum.

OX NCBI_TaxID=3888;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-GV, Mando; TISSUE-Leaf;

EX MEDLINE=94492175; PubMed=4029359;

RA Williamson G.L., Slocum R.D.;

RT "Molecular cloning and characterization of the pybl and pybl2 genes

encoding aspartate transcarbamylase in pea (*Pisum sativum* L.)."

RL Plant Physiol. 105:377-384(1994).

CC -1- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-aspartate -> phosphate

CC -1- ENZYME REGULATION: ALLOSTERICALLY REGULATED BY UMP.

CC -1- PATHWAY: Pyrimidine biosynthesis; second step.

CC -1- SUBUNIT: HOMOTRIMER (POTENTIAL)

CC -1- SIMILARITY: BELONGS TO THE ATCase/OTCase FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/submitter/>

CC or send an email to license@sib-sib.ch).

CC

CC EMBL: M96981; AAA62443.1;

CC HSP: P00479; 3CSU.

CC InterPro: IPR002039; Asp70rn_Corrad;

CC InterPro: IPR002082; Asp_carbamyltransf.

CC Pfam: PF00185; OTCase; 1.

CC Pfam: PF02729; OTCase; N; 1.

CC PRINTS: PP00100; AOTYASF.

CC TRIPRAMS: TRIR00670; asp_carb.tr; 1

CC PROSITE: PS00097; CARBAMOYLTRANSFERASE; 1.

CC Pyrimidine biosynthesis; Transferase; Chloroplast; Transit peptide;

CC Multigene family.

CC

FT TRANSIT

FT CHAIN

SC SEQUENCE 386 AA; 42617 MW; A3440F4574E29F4 CRC64;

Query Match 83.3%; Score 40; DB 1; Length 386;

Best Local Similarity 85.7%; Pred. No. 44;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKEYFTS 7

DB 1 IKEYFTS 281

DB 275 IKEYFTS 281

RESULT 7
 ANCL_PIRSP STANDARD: PRI: 569 AA.
 C P55298;
 T 01-OCT-1996 (Rel. 34, Created)
 T 01-OCT-1996 (Rel. 34, Last sequence update)
 T 15-JUL-1998 (Rel. 36, Last annotation update)
 E Mannan endo-1,4-beta-mannosidase C precursor (EC 3.2.1.78) (beta-mannanase C) (1,4-beta-D-mannan mannanhydrolase C).
 MANC.
 S Piriomyces sp.
 C Eukaryota; Fungi; Chytridiomycota; Neocallimastixales;
 C Neocallimastixaceae; Piriomyces.
 X NCBI_TaxID=45796;
 N [1]
 P SEQUENCE FROM N.A.
 MEDLINE=96313314; PubMed=8768520;
 A Millward-Sadler S.J., Hall J., Black G.W., Hazlewood G.P., Gilbert H.J.;
 T "Evidence that the Piriomyces gene family encoding endo-1,4-mannanases arose through gene duplication."
 L FEMS Microbiol. Lett. 141:183-188(1996).
 C -1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-beta-D-mannosidic linkages in mannans, galactomannans, glucomannans, and galactoglucomannans.
 C -1- SIMILARITY: THE 39 AA REPEATS ARE SIMILAR TO THOSE IN MANC AND XNNA; AND TO THOSE OF N.PATRICIAPUM XNNA.
 C -1- SIMILARITY: BELONGS TO FAMILY 26 OF GLYCOSYL HYDROLASES.
 C This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcements/> or send an email to license@isb-sib.ch).
 C
 R EMBL: X97520; CAA66134.1;
 R InterPro: IPR002883; CBD_5;
 R InterPro: IPR008085; Glyco_hydro_26;
 R InterPro: IPR001230; PreyL_site;
 R Pfam: PF02013; CBM_10; 2;
 R Pfam: PF02156; Glyco_hydro_26; 1;
 R PRINTS: PR00739; GLHYDRLASE26;
 W HydroLase: Glycosidase; Signal; Multigene family; Repeat.
 T SIGNAL 1 18 POTENTIAL.
 T CHAIN 19 569 MANNA ENDO-1,4-BETA-MANNOSIDASE C.
 T DOMAIN 490 569 2 X 39 AA APPROXIMATE REPEATS.
 T REPEAT 488 526 1;
 T REPEAT 531 569 2;
 T DOMAIN 473 477 POLY-ASN.
 T DOMAIN 480 486 POLY-ASN.
 Q SEQUENCE 569 AA; 64115 MW; 19277764E1B328B5 CRC64;
 Query Match 83.38; Score 30; DB 1; Length 569;
 Best Local Similarity At 74; Pred No. 66;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Y 1 IKEYPT 6
 :|||||
 b 444 IKEYPT 449
 :|||||
 RESULT 8
 ANCL_PIRSP STANDARD: PRI: 571 AA.
 C P55297;
 T 01-OCT-1996 (Rel. 34, Created)
 T 01-OCT-1996 (Rel. 34, Last sequence update)
 T 15-JUL-1998 (Rel. 36, Last annotation update)
 E Mannan endo-1,4-beta-mannosidase B precursor (EC 3.2.1.78) (beta-

DE mannanase B) (1,4-beta-D-mannan mannanhydrolase B).
 GN MANB.
 OS Piriomyces sp.
 CC Eukaryota; Fungi; Chytridiomycota; Neocallimastixales;
 CC Neocallimastixaceae; Piriomyces.
 CX NCBI_TaxID=45796;
 RN [1]
 RP SEQUENCE FROM N.A.
 MEDLINE=96313314; PubMed=8768520;
 RA Millward-Sadler S.J., Hall J., Black G.W., Hazlewood G.P., Gilbert H.J.;
 RT "Evidence that the Piriomyces gene family encoding endo-1,4-mannanases arose through gene duplication."
 L FEMS Microbiol. Lett. 141:183-188(1996).
 C -1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-beta-D-mannosidic linkages in mannans, galactomannans, glucomannans, and galactoglucomannans.
 C -1- SIMILARITY: THE 39 AA REPEATS ARE SIMILAR TO THOSE IN MANB, MANC AND XNNA; AND TO THOSE OF N.PATRICIAPUM XNNA.
 C -1- SIMILARITY: BELONGS TO FAMILY 26 OF GLYCOSYL HYDROLASES.
 C This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcements/> or send an email to license@isb-sib.ch).
 C
 DR EMBL: X97408; CAA66061.1;
 DR InterPro: IPR002883; CBD_5;
 DR InterPro: IPR008085; Glyco_hydro_26;
 DR InterPro: IPR001230; PreyL_site;
 DR Pfam: PF02013; CBM_10; 2;
 DR Pfam: PF02156; Glyco_hydro_26; 1;
 DR PRINTS: PR00739; GLHYDRLASE26;
 KW HydroLase: Glycosidase; Signal; Multigene family; Repeat.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 571 MANNA ENDO-1,4-BETA-MANNOSIDASE B.
 FT DOMAIN 490 571 2 X 39 AA APPROXIMATE REPEATS.
 FT REPEAT 490 527 1;
 FT REPEAT 533 571 2;
 FT DOMAIN 475 479 POLY-ASN.
 FT DOMAIN 482 488 POLY-ASN.
 SQ SEQUENCE 571 AA; 64397 MW; B13F44581F8A90AA CRC64;
 Query Match 83.38; Score 40; DB 1; Length 571;
 Best Local Similarity At 74; Pred No. 66;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CY 1 IKEYPT 6
 :|||||
 DB 446 IKEYPT 451
 :|||||
 RESULT 9
 YKAG_YEAST STANDARD: PRI: 245 AA.
 ID YKAG_YEAST
 AC P33201;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical 27.1 kDa protein from YAP1 intergenic region.
 GN YK1009W OR YK1160.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetaceae; Saccharomycetaceae; Saccharomyces.
 CX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 MEDLINE=93127732; PubMed=1481574;
 RX Pascolo S., Glavietz M., Boyer J., Collaenx L., Thierry A., Dujon B.;
 FT "The sequence of a 9.3 kb segment located on the left arm of the

yeast chromosome XI reveals five open reading frames including the CCI1 gene and putative products related to MTO2 and to the ribosomal protein L10.¹¹

[2] MEDLINE-93255900; PubMed-9446424; Boyer J., Paspolo S., Richard G.F., Dujon B.: "Sequence of a 7.8 kb segment on the left arm of yeast chromosome XI reveals four open reading frames, including the CCI1 gene, an intron-containing gene and a gene encoding a homology to the mammalian UOG-1 gene."
Yeast 9:279-287(1993).

-1- SIMILARITY: TO YEAST RIBOSOMAL PROTEIN L10.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch>) or send an email to license@isb-sib.ch).

EMBL: S53418; AAB24904.1; -
EMBL: X61398; -; NOT_ANNOTATED_CDS.
EMBL: Z28009; CAAB1844.1; -
EMBL: S59773; AAC60552.1; -
PIR: S30013; S30013.
SCD: S0001492; YKL009W.
InterPro: IPR001790; Ribosomal_L10.
Pfam: PF00466; Ribosomal_L10. 1.
Hypothetical protein.

SEQUENCE 236 AA. 27058 MW. 3092EEF809F427 CRCK4;

Query Match 80.6%; Score 29; DB 1; Length 236;
Best Local Similarity 71.4%; Pred. No. 42;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 IKEYPTS 7
:||||:
112 VKRYFS 118

CA_THEME STANDARD; PRT: 338 AA.
Q9WZV5: 30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Heat-inducible transcription repressor HrcA homolog
HRC A OR TM0851.
Thermotoga maritima
Bacteria; Thermotogae; Thermotogae (class); Thermotogales.
Thermotogaceae; Thermotoga.
NCBI_TaxID=2336;

[1] SEQUENCE FROM N.A.
STRAIN-MSB8 / DSM 3109;
MEDLINE-99287316; PubMed-10360571;
Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J., Hall D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A., McDonald L., Utterback T.P., Majek J.A., Linher K.D., Garrett M.M., Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D., Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O., Salzberg S.L., Smith H.C., Venter J.C., Fraser C.M.:
"Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of Thermotoga maritima."
Nature 399:323-329(1999).

-1- FUNCTION: NEGATIVE REGULATOR OF CLASS I HEAT SHOCK GENES (GPEE-DNAK-DNAJ AND GPEELS OPERONS). PREVENTS HEAT-SHOCK INDUCTION OF THESE OPERONS (BY SIMILARITY).

-1- SIMILARITY: BELONGS TO THE HRC A FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch>)
CC or send an email to license@isb-sib.ch).

EMBL: AE001751; AAD35933.1; -
DR TIGR: TM0851; -
DR InterPro: IPR002571; HrcA.
PFam: PF01628; HrcA; 1.
DE TIGRFAMs: TIGR00331; hrcA; 1.
KW Transcription regulation; Repressor; Heat shock; Complete proteome.
SQ SEQUENCE 338 AA. 39306 MW. 02A05A546731AE CRCK4;

Query Match 80.6%; Score 29; DB 1; Length 338;
Best Local Similarity 71.4%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 IKEYPTS 7
:||||:
328 LSEYFS 334

RESULT 11
ACBC_ACT55 STANDARD; PRT: 342 AA.
ID ACBC_ACT55
AC Q9ZAE9; 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
RT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 2-epi-5-epi-valiolone synthase (EC 4.2.3.-) (Sedo-heptulose 7-
phosphate cyclase).
GN ACBC.
OS Actinoplanes sp. (strain 50/110).
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteriales;
OC Actinomycetales; Micromonosporineae; Micromonosporaceae;
OC Actinoplanes.
OX NCBI_TaxID=1871;
RN [1]

SEQUENCE FROM N.A.
MEDLINE-99214159; PubMed-10196166;
RA Strittman A., Mahmud T., Lee S., Distler J., Floss H.G.,
RA Piepersberg W.;
RT "The AcBC protein from Actinoplanes species is a C7-cyclitol synthase
related to 3-dehydroquinate synthases and is involved in the
biosynthesis of the alpha-glucosidase inhibitor arabinose."
RL J. Biol. Chem. 274:10889-10896(1999).
CC -1- FUNCTION: C7-cyclitol synthase using sedo-heptulose 7-phosphate,
CC but not ido-heptulose 7-phosphate and 3-deoxy-arabino-
CC heptulose 7-phosphate, as a substrate. Involved in the
CC biosynthesis of the acarbose moiety of the alpha-glucosidase
CC inhibitor acarbose.
CC -1- CATALYTIC ACTIVITY: Sedo-heptulose 7-phosphate + 2-epi-5-epi-
CC -1- VALIOLONE + phosphate.
CC -1- COFACTOR: NAD and a divalent metal cation such as cobalt.
CC -1- SIMILARITY: BELONGS TO THE DEHYDROQUINATE SYNTHASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch>)
CC or send an email to license@isb-sib.ch).

EMBL: Y18523; CAA77208.1; -
DR InterPro: IPR002458; DHQ_synthase.
DR Pfam: PF01761; DHQ_synthase; 1.
KW Lyase; NAD; Cobalt.
SQ SEQUENCE 382 AA. 42181 MW. EDA70004A144B447 CRCK4;

Query Match 80.6% Score 29, DB 1, Length 382;
Best Local Similarity 83.3% Pred. No. 71;
Matches 5: Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYFT 6
|||||
73 REEFT 78

RESULT 12
SD_BACSU
D CSD_BACSU STANDARD: PRT: 406 AA.

032164;
16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
Probable cysteine desulfurase (EC 4.4.1.-).
CSD.
Bacillus subtilis.
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_TaxID:1423;
[1]
SEQUENCE FROM N.A.
STRAIN:168;
MEDLINE:98044033; PubMed:9184377;
Azevedo V., Oliveira N., Moszer I., Albertini A.M., Alloni G.,
Boriss R., Boursier L., Brans A., Braun M., Brigelli S.C., Bron S.,
Broillet S., Hirsch C.V., Caldwell B., Capano V., Carter N.M.,
Choi S.K., Colaianni I., Cornet I.F., Cummings N.C., Daniel R.A.,
DeLutz P., Levine K.M., Dusterhoft A., Enlich S.D., Emerson P.T.,
Enlian K.D., Erismann J., Fabre C., Feriail E., Foulger L.,
Fritz G., Glaser M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
Ghim S.Y., Glaeser P., Goffeau A., Gollightly E.J., Grandi G.,
Guspehl G., Guy R.J., Hara K., Harlow J., Harwood C.P., Haurat A.,
Hilbert H., Holsappel S., Hosono S., Huijo M.F., Itaya M., Jones L.,
Joris B., Karamata D., Kashara Y., Kiebert-Blochard M., Klein C.,
Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
Kurita K., Lapidis A., Lardinois S., Lauber J., Lazarevic V.,
Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medina C.,
Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
Noone D., O'Leary M., Ogawa K., Ojima A., Oudega R., Park S.H.,
Pardo V., Pohl T.M., Portetelle D., Potworlik S., Prescott A.M.,
Presecan E., Puig P., Puttelle R., Rapoport G., Pey M., Reynolds S.,
Rivet M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
Sekiguchi I., Sekowska A., Seror S.J., Seirer P., Shin B.S., Solde B.,
Sorokin A., Tachikawa T., Takai T., Takahashi H., Takemura Y.,
Takouchi M., Tamakoshi A., Tanaka T., Terpestra P., Tognoni A.,
Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassaretzi A.,
Vialti A., Wambur P., Wadler P., Waller H., Wellenreger T.,
Winzler E., Wipat A., Yamamoto H., Yamada K., Yasunaga K., Yata K.,
Yoshida K., Yoshikawa H.F., Ziemstein E., Yoshikawa H., Bachlin A.,
The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*;
L Nature 390:249-256(1997).
- I- FUNCTION: CATALYZES THE REMOVAL OF ELEMENTAL SULFUR AND SELENIUM
ATOMS FROM L-CYSTEINE, L-CYSTEINE, L-SELENOCYSTEINE, AND L-
SELENOCYSTEINE TO PRODUCE L-ALANINE (BY SIMILARITY).
- I- CATABOLIC: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
- I- SIMILARITY: BELONGS TO CLASS-V OF PYRIDOXAL-PHOSPHATE-DEPENDENT
AMINOTRANSFERASES CSD SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed, usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/annotation/>
or send an email to license@isb-sib.ch).

EMBL: Z99120; CAB15258.1; -

DR HSP: P77444; 1JF9;
DR Subtilisin; B014009; Csd.
DR InterPro: IPR000192; Aminotransf.
DR Pfam: PF00266; aminotran5; 1;
DR PROSITE: PS00595; AA-TRANSFER CLAS; 5; 1;
KW lysine, pyridoxal phosphate, complete, proteome,
PT BINDING 224
PT BINDING 224
SQ SEQUENCE 406 AA; 44922 MW; E4300B595349D1 CRC64;

Query Match 80.6% Score 29, DB 1, Length 406;
Best Local Similarity 83.3% Pred. No. 75;
Matches 5: Conservative 1; Mismatches 0; Indels 0; Gaps 0;

2 KEYFTS 7
|||||
399 KEYFTN 404

RESULT 13
YSDI_CAEEL
ID YSDI_CAEEL STANDARD: PRT: 445 AA.

AC 010127;
01-FEB-1996 (Rel. 33, Created)
01-FEB-1996 (Rel. 33, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical zinc finger protein P00111; in chromosome 11;
GN F56D1.1;
ON Caenorhabditis elegans.
NC F047424; M047424; N047424; Chrom14:79; Phalloidin; Phalloidin;
NC Phalloidin; Pelodermidae; Caenorhabditis.
NCBI_TaxID:6216;
[1]
SEQUENCE FROM N.A.
STRAIN:Bristol N2;
RA Chinoe S., Wilson R.J.
FL Submitted (Nov 1995) to the EMBL/GenBank/CCRF databases.
- I- SUBCELLULAR LOCATION: Nuclear (potential).

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed, usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/annotation/>
or send an email to license@isb-sib.ch).

EMBL: U39997; AAA81101.1; -
Wormpep: F56D1.1; CE01969.
01-FEB-1996 (Rel. 33, Created)
01-FEB-1996 (Rel. 33, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein; Zinc finger; Zinc finger; Zinc finger;
KW Hypothetical protein; Zinc finger; Zinc finger; Zinc finger;
KW Nuclear protein; Repeat.
FT 2N-FING 46
FT 2N-FING 68
FT 2N-FING 98
FT 2N-FING 103
FT 2N-FING 130
FT 2N-FING 137
FT 2N-FING 317
FT 2N-FING 340
FT 2N-FING 419
SQ SEQUENCE 435 AA; 50097 MW; 240CA095F75343D6 CRC64;

Query Match 80.6% Score 29, DB 1, Length 435;
Best Local Similarity 85.7% Pred. No. 81;
Matches 6: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 KEYFTS 7
|||||
150 KEYFTS 156

SOUT 14
 RC_CLOPE STANDARD: PRI: 620 AA.
 08XN11:
 15-JUN-2002 (Rel. 41, Created)
 15-JUN-2002 (Rel. 41, Last sequence update)
 15-JUN-2002 (Rel. 41, Last annotation update)
 Exonuclease ABC subunit C.
 UVRC OR CPE0352.
 Clostridium perfringens.
 Bacteria: Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 Clostridium
 NCBI_TaxID=1502;
 [1]
 SEQUENCE FROM N.A.
 STRAIN-13 / Type A;
 PubMed=11792842;
 Shimizu T., Ohnami K., Hirakawa H., Ohshima K., Yamashita A.,
 Shibata T., Ogasawara N., Hattori M., Kubara S., Hayashi H.;
 *Complete genome sequence of Clostridium perfringens, an anaerobic
 flesh-eater.*;
 Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
 -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
 CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
 PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). ATTACHES TO THE
 UVRA-UVRC COMPLEX, DISPLACING UVRA, AND THE DAMAGED DNA STRAND IS
 NICKED ON BOTH SIDES OF THE DAMAGED SITE (BY SIMILARITY).
 -1- SUBUNIT: Consists of three subunits: UVRA, UVRB and UVRC.
 -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 -1- SIMILARITY: BELONGS TO THE UVRC FAMILY.
 -1- SIMILARITY: CONTAINS 1 TYP DOMAIN.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb.slb.ch/announce/>
 or send an email to license@isb.slb.ch).

 EMBL: AP003186; BAB80058.1;
 InterPro: IPR000445; HhH.
 InterPro: IPR001943; UVRC.
 InterPro: IPR004791; UVRC.
 InterPro: IPR001162; UVRC.
 InterPro: IPR000305; UVRC_N.
 Pfam: PF01541; Exclendo_N.1.
 Pfam: PF02151; UVRC.1.
 ProDom: PD005870; UVRC_C.1.
 SMART: SM00465; GYRC.1.
 TIGRFAMs: TIGR00194; uvrc.1.
 PROSITE: PS0151; UVRC.1
 SOS response; Excision nuclease; DNA repair; Complete proteome.
 DOMAIN 204 239 UVRC
 SEQUENCE 620 AA; 71390 MW; 6E14ADE9C3849219 CRC64;
 Query Match 80.6%; Score 29; DB 1; Length 620;
 Best Local Similarity 71.4%; Pred. No. 1.2e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 1 KEYFITS 7
 1111111
 614 IKOYESS 620
 SOUT 15
 81_YEAST STANDARD: PRT: 1562 AA.
 004781: 004029;
 01-NOV-1987 (Rel. 35, Created)
 01-NOV-1987 (Rel. 35, Last sequence update)
 16-OCT-2001 (Rel. 40, Last annotation update)
 Hypothetical 180.2 kDa protein in FAA4-HOR7 intergenic region.

GN YMR247C OR YMR408 ORC OR YMR920 ORC
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 FN [1]
 PP SEQUENCE OF 1-956 FROM N.A.
 RC STRAIN-S288c / AB972;
 RA Hunt S., Bowman S., Barrell B.G., Rajandream M.A., Walsh S.V.;
 RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 950-1562 FROM N.A.
 RC STRAIN-S288c / AB972;
 RA Gentles S., Bowman S., Barrell B.G., Rajandream M.A., Walsh S.V.;
 RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 CC -1- SIMILARITY: STRONG, TO HUMAN ZNF744.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb.slb.ch/announce/>
 or send an email to license@isb.slb.ch).
 CC -----
 DR EMBL: Z48639; CAAB8574.1;
 DR EMBL: Z48756; CAAB8657.1;
 DR S50: S0004861; YMR247C.
 DR InterPro: IPR001841; ZnF_Ring.
 DR SMART: SM00184; RING.1.
 DR PROSITE: PS00518; ZF_RING_1; FAISE_NHG.
 DR PROSITE: PS00893; ZF_RING_2; 1.
 KW Hypothetical protein; Zinc-finger.
 FT ZN-FINGER 1508 1555 RING-TYPE
 SQ SEQUENCE 1562 AA; 180185 MW; 97AC55E81362305 CRC64;
 Query Match 80.6%; Score 29; DB 1; Length 1562;
 Best Local Similarity 71.4%; Pred. No. 3.1e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 1 KEYFITS 7
 1111111
 Db 462 LEDYFTS 468
 Search completed: April 8, 2003, 12:02:17
 Job time : 19.9167 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

M protein - protein search, using sw model

Search time 30.9167 seconds

(without alignments)
21.766 Million cell updates/sec

Sequence: 1 IKEFITS 7

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Number of hits satisfying chosen parameters: 28324

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

2: F12.*
3: P13.*
4: P14.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	88.9	279	2 E72425	dihydropterate synthase
2	32	88.9	406	2 T32922	hypothetical prote
3	32	88.9	416	2 S44411	hypothetical prote
4	32	88.9	469	2 S55167	hypothetical prote
5	32	88.9	577	2 A57469	hypothetical prote
6	32	88.9	620	2 E90575	hypothetical prote
7	32	88.9	3562	2 A47171	hypothetical prote
8	31	86.1	260	2 T50109	hypothetical prote
9	31	86.1	263	2 T51235	hypothetical prote
10	31	86.1	1460	2 S48457	hypothetical prote
11	30	83.3	150	2 S55903	hypothetical prote
12	30	83.3	150	2 E64376	hypothetical prote
13	30	83.3	150	2 A90991	hypothetical prote
14	30	83.3	150	2 C85836	hypothetical prote
15	30	83.3	153	2 S47625	hypothetical prote
16	30	83.3	386	2 T06484	hypothetical prote
17	30	83.3	402	2 H86649	hypothetical prote
18	30	83.3	487	2 T27353	hypothetical prote
19	30	83.3	524	2 T44988	hypothetical prote
20	30	83.3	524	2 C07703	hypothetical prote
21	30	83.3	573	2 T05338	hypothetical prote
22	30	83.3	598	2 T93824	hypothetical prote
23	30	83.3	1090	2 H86806	hypothetical prote
24	29	80.6	92	2 A62174	hypothetical prote
25	29	80.6	136	2 C71661	hypothetical prote
26	29	80.6	176	2 F01800	hypothetical prote
27	29	80.6	199	2 E99972	hypothetical prote
28	29	80.6	214	2 H97767	hypothetical prote
29	29	80.6	240	2 T01601	hypothetical prote

ALIGNMENTS

RESULT 1

E72425

dihydropterate synthase (EC 2.5.1.15) [tm004] [similarity] - Thermotoga maritima (str

C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence-revision 11-Jun-1999 #text-change 02-Sep-2000

C:Accession: E72425

R:Nelson, K.E.; Clayton, R.A.; Gill, S.P.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, C.M.

Nature 359, 323-329, 1999

Article: Evidence for lateral gene transfer between Archaea and Bacteria from genome

A:Reference number: A72000 M010424700 TMDP:0440071

A:Accession: E72425

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-278 <ARN>

A:Cross-references: GB:AF001691; GB:AF005512; NID:94960517; FIDN:AND3134.1; PID:9498

A:Experimental source: strain MSB8

C:Genetics:

A:Superfamily: dihydropterate synthase, dihydropterate synthase homology

C:Keywords: folate biosynthesis; transforase

E:18-58/Domain: dihydropterate synthase homology -CBS>

Query Match

Best Local Similarity 88.9% Score 12; DB 2; Length 278;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 IKEFIT 6

DB 163 IKEFIT 168

RESULT 2

T32922

hypothetical protein K09H9.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence-revision 29-Oct-1999 #text-change 20-Oct-1999

C:Accession: T32922

R:Madson, C.; Graves, T.; Blatt, T.

Submitted to the EMBL Data Library, January 1998

A:Reference number: T32922

A:Accession: T32922

A>Status: preliminary; translated from cDNA

A:Molecule type: DNA

A:Residues: 1-406 <MAT>

A:Cross-references: EMBL:AF047000; FIDN:ZAR7522.1; GSTRB:3800019; CEST:K09H9.5

A:Experimental source: strain Bristol N. clone K09H9

C:Genetics:

A:Gene: CESP-K09H9.5

A:Map position: 1

Reference number: S55901; MUID:95290497; PMID:7772602
 Accession: S55903
 Status: Preliminary
 Molecule type: DNA
 Residues: 1-150 <NR>
 Cross-references: EMBL:X78837, NID:9579737, FIDN:GMS6228.1, PID:9509173
 Genes:
 Gene: gata

Query Match 83.3% Score 30; DB 2; Length 150;
 Best Local Similarity 83.3%; Pred. No. 41;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYFT 6
 :|||||
 137 LKEYFT 142

SUIT 12
 4976
 s system, galactitol-specific IIA component - Escherichia coli (strain K-12)

Species: Escherichia coli
 Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
 Accession: E64976

Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co-
 .. Rose, D.J.; Mau, B.; Shaoy, Y.

Title: The complete genome sequence of Escherichia coli K-12.
 Reference number: A64720; MUID:97426617; PMID:9278503

Accession: E64976
 Status: Preliminary; nucleic acid sequence not shown; translation not shown

Molecule type: DNA
 Residues: 1-150 <BLAT>
 Cross-references: GB:AE000294; GB:066096; NID:91786195; FIDN:AGC5155.1; PID:91788410.

Experimental source: strain K-12, substrain MG1655
 Genetics:

Query Match 83.3% Score 30; DB 2; Length 150;
 Best Local Similarity 83.3%; Pred. No. 41;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYFT 6
 :|||||
 137 LKEYFT 142

SUIT 13
 0991
 s system, galactitol-specific IIA component - Escherichia coli (strain 0157:H7, substr

Species: Escherichia coli
 Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 27-Nov-2001
 Accession: A90991

Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Sawara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
 Reference number: A99629; MUID:21156231; PMID:11258796

Accession: A90991
 Status: Preliminary

Molecule type: DNA
 Residues: 1-150 <HAV>
 Cross-references: GB:BA000007; PIDN:BA836720.1; PID:91362466; GSFD:GN00154

Experimental source: strain 0157:H7, substrain RMD 0509952
 Genetics:

Gene: EGS2897

Query Match 83.3% Score 30; DB 2; Length 150;
 Best Local Similarity 83.3%; Pred. No. 41;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYFT 6
 :|||||

Db 137 LKEYFT 142

RESULT 14

C85836
 s system, galactitol-specific IIA component - Escherichia coli (strain 0157:H7, sub

Species: Escherichia coli
 Date: 16-Feb-2001 #sequence_revision 17-Feb-2001 #text_change 27-Nov-2001
 Accession: C85836

R. Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
 Hille, L.; Grothbeck, F.T.; Davis, N.W.; Lim, A.; Diallanla, E.; Potamotis, K.; Apoda

Nature 409, 529-533, 2001
 A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: C85836
 A:Status: Preliminary

A:Molecule type: DNA
 A:Residues: 1-150 <STO>

A:Cross-references: GB:AE005174; NID:91362297; PIDN:AGC5151.1; GSFD:GN00145; UWGP:
 A:Experimental source: strain 0157:H7, substrain EDL933
 A:Genetics:

A:Gene: gata

Query Match 83.3% Score 30; DB 2; Length 150;
 Best Local Similarity 83.3%; Pred. No. 41;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYFT 6
 :|||||
 137 LKEYFT 142

RESULT 15
 S47625
 aspartate carbamoyltransferase (EC 2.1.1.4) - wheat (fragment)

Alternate names: aspartate transcarbamoylase
 C:Species: Triticum aestivum (common wheat)
 C:Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 07-May-1999

C:Accession: S47625
 R:Bartlett, T.J.; Albangbee, A.; Bruce, I.J.; Donovan, P.J.; Yon, R.J.

Biotech. Biochem. Acta 1207, 187-193, 1994
 A>Title: Endogenous polypeptide-chain length and partial sequence of aspartate trans

A:Reference number: S47625; MUID:94355374; PMID:8075153

A:Accession: S47625
 A:Status: not compared with conceptual translation

A:Molecule type: mRNA
 A:Residues: 1-233 <BAR>

A:Experimental source: var. Avalon
 C:Superfamily: ornithine carbamoyltransferase; aspartate/ornithine carbamoyltransferase
 C:Keywords: pyrimidine nucleotide biosynthesis; transferase; homotrimer

E1-223; domain: aspartate/ornithine carbamoyltransferase homology (fragment) <ACT>

Query Match 83.3% Score 30; DB 2; Length 233;
 Best Local Similarity 85.7%; Pred. No. 64;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 KEYFTS 7
 :|||||
 Db 115 IKEYFTS 121

Search completed: April 8, 2003, 12:05:13
 Job time: 34.9167 secs

GenCore version 5.1.4_P5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

M protein - protein search, using sw model

un 00: April 8, 2003, 12:01:50 : Search time 12.8333 seconds

(without alignments)
33.347 Million cell updates/sec

sequence: 1 KEFFTS 7

scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

searched: 248812 seqs, 61136040 residues

total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match Length	DB ID	Description
1	30	83.3	150 10 US-09-741-669-424	Sequence 424, App
2	29	80.7	6745 10 US-09-815-242-1518	Sequence 12610, A
3	28	77.8	141 10 US-09-815-242-11289	Sequence 11471, A
4	28	77.8	141 10 US-09-815-242-11471	Sequence 11471, A
5	28	77.8	219 10 US-09-782-874-7	Sequence 3, Appl
6	28	77.8	224 10 US-09-012-020-359	Sequence 350, App
7	28	77.8	276 12 US-10-033-078-16	Sequence 16, Appl
8	28	77.8	377 10 US-09-910-430-27	Sequence 27, Appl
9	28	77.8	431 9 US-10-081-816-5	Sequence 5, Appl
10	28	77.8	511 9 US-09-734-026-5406	Sequence 5406, AP
11	28	77.8	573 10 US-09-815-242-10760	Sequence 10760, A
12	28	77.8	584 10 US-09-770-940-2	Sequence 2, Appl
13	28	77.8	724 10 US-09-815-242-6093	Sequence 6093, AP
14	28	77.8	1114 10 US-09-782-874-2	Sequence 2, Appl
15	27	75.0	33 10 US-09-144-838-14	Sequence 14, Appl
16	27	75.0	33 10 US-09-144-838-13	Sequence 13, Appl
17	27	75.0	33 10 US-09-144-838-15	Sequence 15, Appl
18	27	75.0	34 10 US-09-144-838-16	Sequence 16, Appl
19	27	75.0	45 12 US-10-001-870-147	Sequence 147, Mf

20	27	75.0	53 9 US-10-014-452-12	Sequence 12, Appl
21	27	75.0	60 9 US-09-888-938-5	Sequence 5, Appl
22	27	75.0	62 9 US-09-809-391-486	Sequence 486, App
23	27	75.0	66 10 US-09-144-838-35	Sequence 35, Appl
24	27	75.0	66 10 US-09-144-838-37	Sequence 37, Appl
25	27	75.0	67 10 US-09-144-838-31	Sequence 31, Appl
26	27	75.0	67 10 US-09-144-838-33	Sequence 33, Appl
27	27	75.0	67 10 US-09-144-838-36	Sequence 36, Appl
28	27	75.0	67 10 US-09-144-838-38	Sequence 38, Appl
29	27	75.0	67 10 US-09-144-838-39	Sequence 39, Appl
30	27	75.0	67 10 US-09-144-838-41	Sequence 41, Appl
31	27	75.0	68 9 US-09-742-794-29	Sequence 29, Appl
32	27	75.0	68 10 US-09-864-761-40180	Sequence 40180, A
33	27	75.0	68 10 US-09-144-838-19	Sequence 19, Appl
34	27	75.0	68 10 US-09-144-838-32	Sequence 32, Appl
35	27	75.0	68 10 US-09-144-838-34	Sequence 34, Appl
36	27	75.0	68 10 US-09-144-838-40	Sequence 40, Appl
37	27	75.0	68 10 US-09-144-838-42	Sequence 42, Appl
38	27	75.0	68 10 US-09-144-838-43	Sequence 43, Appl
39	27	75.0	68 10 US-09-144-838-45	Sequence 45, Appl
40	27	75.0	68 10 US-09-195-457-11	Sequence 11, Appl
41	27	75.0	69 10 US-09-144-838-44	Sequence 44, Appl
42	27	75.0	69 10 US-09-144-838-46	Sequence 46, Appl
43	27	75.0	89 10 US-09-144-477-8	Sequence 4, Appl
44	27	75.0	92 8 US-09-742-794-21	Sequence 21, Appl
45	27	75.0	91 9 US-10-017-275-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-09-741-669-424
Sequence 424, Application: 15, 0017069
FASTA ID: US-09-741-669-424
GENERAL INFORMATION:
APPLICANT: Forsyth, P. Allyn
APPLICANT: Obisen, Karl L.
TITLE OR INVENTION: Genes identified as required for
FILE REFERENCE: Elifira 0094
CURRENT APPLICATION NUMBER: US 01/74 449
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 60/370055
PRIOR FILING DATE: 1999-12-23
NUMBER OF SEQ ID NOS: 481
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 424
LENGTH: 150
TYPE: PRI
ORGANISM: Escherichia coli
US-09-741-669-424

Query Match: 83.3% Score 30: DB ID: Length 150:
Best local similarity at 78: Pred. NO: 61:
Matches: 5, conservative: 1, M: matches: 0, Indels: 0, Gaps: 0:

27 1 KEFFTS 6
DB 127 KEFFTS 142

RESULT 2
US-09-815-242-12610
Sequence 12610, Appl: 157156 US-09-815-242-12610
PATENT NO: US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Obisen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Travis, John D.

APPLICANT: Carr, Grant J.
 APPLICANT: Yamamoto, Robert T.
 TITLE OF INVENTION: Identification of Essential Genes in
 FILE REFERENCE: ELITRA.011A
 CURRENT APPLICATION NUMBER: US/09/815,242
 CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/191,074
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 14110
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 12610
 LENGTH: 5795
 TYPE: PRT
 ORGANISM: *Staphylococcus aureus*
 -09-815-242-112610
 Query Match 80.6% Score 29 DB 10 Length 5795
 Best Local Similarity 83.3% Pred. No. 38703
 Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

1 KEYFT 6
 111111
 1127 IKYFT 1152

SULT 3
 -09-815-242-11269
 Sequence 11269, Application US/9815242
 Patent No. US2002006159A1
 GENERAL INFORMATION:
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Karl L.
 APPLICANT: Zyskind, Judith W.
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John D.
 APPLICANT: Carr, Grant J.
 APPLICANT: Yamamoto, Robert T.
 APPLICANT: Xu, H. Howard
 TITLE OF INVENTION: Identification of Essential Genes in
 FILE REFERENCE: ELITRA.011A
 CURRENT APPLICATION NUMBER: US/09/815,242
 CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 14110
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 14110
 LENGTH: 5795
 TYPE: PRT
 ORGANISM: *Helicobacter pylori*
 -09-815-242-11471

SEQ ID NO 11289
 LENGTH: 141
 TYPE: PRT
 ORGANISM: *Helicobacter pylori*
 US-09-815-242-11289
 Query Match 77.8% Score 24 DB 10 Length 141
 Best Local Similarity 100.0% Pred. No. 1148702
 Matches 5: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

2 KEYFT 6
 111111
 71 KEYFT 75

RESULT 4
 US-09-815-242-11471
 Sequence 11471, Application US/09815242
 Patent No. US2002006159A1
 GENERAL INFORMATION:
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Karl L.
 APPLICANT: Zyskind, Judith W.
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John D.
 APPLICANT: Carr, Grant J.
 APPLICANT: Yamamoto, Robert T.
 APPLICANT: Xu, H. Howard
 TITLE OF INVENTION: Identification of Essential Genes in
 FILE REFERENCE: ELITRA.011A
 CURRENT APPLICATION NUMBER: US/09/815,242
 CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 14110
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 11471
 LENGTH: 141
 TYPE: PRT
 ORGANISM: *Helicobacter pylori*
 US-09-815-242-11471

Query Match 77.8% Score 24 DB 10 Length 141
 Best Local Similarity 100.0% Pred. No. 1148702
 Matches 5: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

2 KEYFT 6
 111111
 71 KEYFT 75

RESULT 5
 US-09-782-874-3
 Sequence 3, Application US/09782874
 Patent No. US20010023067A1
 GENERAL INFORMATION:
 APPLICANT: Wassenecker, Michael
 APPLICANT: Riedel, Leonhard
 APPLICANT: Schiebel, Winfried
 APPLICANT: Sanger, Heinz

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
POLYPEPTIDES HAVING THE ENZYMATIC ACTIVITY OF AN
RNA-DIRECTED RNA POLYMERASE (RDRP)

NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 09/647,749
FILING DATE: 09-Feb-2001
CLASSIFICATION: <UNKNOWN>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/411,583
FILING DATE: 05-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Haley, James F.
REGISTRATION NUMBER: 27,794
REFERENCE/PATENT NUMBER: M9-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 218 amino acids
TYPE: amino acid
STRANDS: 1
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-782-874-3

Query Match 77.8% Score 28; DB 10; Length 218;
Best Local Similarity 57.1%; Pred No. 2.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

1 KEYFITS 7
DB 144 VEEFFTS 150

RESULT 6
US-09-912-020-350
Sequence 350; Application US/09/012000
Patent No. US2002004592A1
GENERAL INFORMATION:
APPLICANT: Zyskind, Judith
APPLICANT: Ohlsen, Karl L
APPLICANT: Trawick, John
APPLICANT: Forsyth, R. Allan
APPLICANT: Froelich, Jamie M.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: GENTS IDENTIFIABLE AS PEPTIDES FOR DIFFERENTIATION IN
TITLE OF INVENTION: ESCHERICHIA COLI
FILE REFERENCE: ELITPA 00101
CURRENT APPLICATION NUMBER: US 09/912,000
CURRENT FILING DATE: 2001-07-23
PRIOR APPLICATION NUMBER: 09/432,706
PRIOR FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: 09/117,405
PRIOR FILING DATE: 1999-01-27
NUMBER OF SEQ ID NOS: 485
SOFTWARE: FASTSEQ for Windows Version 3.0

SEQ ID NO 350
LENGTH: 234
TYPE: PRT
ORGANISM: E. COLI
US-09-912-020-350

Query Match 77.8% Score 28; DB 10; Length 234;
Best Local Similarity 63.3%; Pred No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYFITS 6
DB 224 IKEYFFTS 229

RESULT 7
US-10-033-078-16
Sequence 16; Application US/10/033078
Patent No. US2002012109A1
GENERAL INFORMATION:
APPLICANT: Furukuchi, Yasuhiko
APPLICANT: Hoshino, Tatsuo
APPLICANT: Kimura, Hiroshi
APPLICANT: Kiyasu, Tatsuya
APPLICANT: Nagahashi, Yoshie
TITLE OF INVENTION: RIGGIN BIOSYNTHETIC GENES
FILE REFERENCE: Biotin Genes
CURRENT APPLICATION NUMBER: US/10/033078
CURRENT FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 09/594,184
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: EP 961155A0.5
PRIOR FILING DATE: 1996-09-27
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver 2.1
SEQ ID NO 16
LENGTH: 276
TYPE: PRT
ORGANISM: Kurthia sp.
US-10-033-078-16

Query Match 77.8% Score 28; DB 10; Length 276;
Best Local Similarity 57.1%; Pred No. 2.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

1 KEYFITS 7
DB 214 VEEFFTS 220

RESULT 8
US-09-913-430-27
Sequence 27; Application US/09/910430
Patent No. US2002012735A1
GENERAL INFORMATION:
APPLICANT: Godtrot, Edmond
APPLICANT: Bollen, Alex
APPLICANT: Lebouille, Gerard
TITLE OF INVENTION: IDENTIFICATION AND MOLECULAR CHARACTERIZATION OF
TITLE OF INVENTION: GLANDS, EXPRESSED IN THE XODUS RICHINUS SALIVARY
FILE REFERENCE: VANM29.001CPI
CURRENT APPLICATION NUMBER: US 09/913,430
CURRENT FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: PCT/FR99/0041
PRIOR FILING DATE: 1999-06-06
PRIOR APPLICATION NUMBER: GB 991347.5
PRIOR FILING DATE: 1999-06-09
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 27
LENGTH: 377
TYPE: PRT

ORGANISM: Ixodes ricinus
US-09-910-430-27

Query Match 77.8% Score 28; DB 10; Length 377;
Best Local Similarity 57.18; Pred. No. 3.6e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

1 KEYFT 7
|||||
290 VKDFTS 296

RESULT 9
US-10-081-816-5
Sequence 5; Application US/0908165
Publication No. US2003045472A1

GENERAL INFORMATION:

APPLICANT: Axel Richard

TITLE OF INVENTION: Chemosensory Gene Family Encoding Gustatory And Olfactory Receptors

FILE REFERENCE: 0575-6414-A/JP/AEM

CURRENT APPLICATION NUMBER: US/10/081.816

PRIOR FILING DATE: 2002-02-22

PRIOR APPLICATION NUMBER: 60/271,319

NUMBER OF SEQ ID NOS: 116

SOFTWARE: PatentIn version 3.1

SEQ ID NO 5

LENGTH: 431

TYPE: PRT

ORGANISM: Drosophila melanogaster

US-10-081-816-5

Query Match 77.8% Score 28; DB 9; Length 421;
Best Local Similarity 83.38; Pred. No. 4.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYFT 6
|||||

286 IEVEFT 291

RESULT 10
US-09-748-626-5606
Sequence 5606; Application US/09738626
Publication No. US20020197605A1

GENERAL INFORMATION:

APPLICANT: NAKASAMA, SATOSHI

APPLICANT: MIZOGUCHI, HIROSHI

APPLICANT: ANDO, SEIKO

APPLICANT: HAYASHI, MIKIO

APPLICANT: OCHIAI, KEIKO

APPLICANT: YOKOI, HARUHIKO

APPLICANT: TATEISHI, NACKO

APPLICANT: SENOH, AKIHIRO

APPLICANT: IKEDA, MASATO

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-125

CURRENT APPLICATION NUMBER: US/09/748.626

PRIOR FILING DATE: 2000-12-18

PRIOR APPLICATION NUMBER: JP 99/377484

PRIOR FILING DATE: 1999-12-18

PRIOR APPLICATION NUMBER: JP 00/159162

PRIOR FILING DATE: 2000-04-07

PRIOR APPLICATION NUMBER: JP 00/280988

PRIOR FILING DATE: 2000-08-03

NUMBER OF SEQ ID NOS: 7059

SOFTWARE: PatentIn ver. 3.0

SEQ ID NO 5606

LENGTH: 511

TYPE: PRT

ORGANISM: Corynebacterium glutamicum
US-09-738-626-5606

Query Match 77.8% Score 28; DB 9; Length 511;
Best Local Similarity 83.38; Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYFT 6
|||||
386 IEVEFT 391

RESULT 11
US-09-815-242-10760
Sequence 10760; Application US/09815242
Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl D.

APPLICANT: Tysland, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant T.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: FLITRA.01A

CURRENT APPLICATION NUMBER: US/09/815.242

PRIOR FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

FILE REFERENCE: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 10760

LENGTH: 570

TYPE: PRT

ORGANISM: Enterococcus faecalis

US-09-815-242-10760

Query Match 77.8% Score 28; DB 10; Length 570;
Best Local Similarity 100.08; Pred. No. 5.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 KEYFT 6
|||||
247 KEVEFT 251

RESULT 12

US-09-770-940-2
Sequence 2; Application US/09770940
Patent No. US20020009790A1

GENERAL INFORMATION:

APPLICANT: Christensen, T.

APPLICANT: Thorsen, H.

APPLICANT: Kreider, J.

APPLICANT: Buchholt, H.

TITLE OF INVENTION: PROCESS FOR STABILIZING PROTEINS IN AN

FILE REFERENCE: DYO014.001APC

```
CURRENT APPLICATION NUMBER: US/09/770,940
CURRENT FILING DATE: 2001-01-24
EARLIER APPLICATION NUMBER: 08/983,364
EARLIER FILING DATE: 1998-05-18
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO: 2
LENGTH: 584
TYPE: PRT
ORGANISM: Citrus hawaiiensis var. class 1
US-09-770-940-2

Query Match: 77.8% Score 28; DB 10; Length 584;
Best Local Similarity 57.1%; Pred. No. 5.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Y 1 KEYPTS 7
  1:1111
Db 4 KEYPTS 9

RESULT 13
US-09-815-242-5093
Sequence 5093: Application US/09/815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Ohlssen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITPA.01A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO: 5093
LENGTH: 723
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5093

Query Match: 77.8% Score 28; DB 10; Length 723;
Best Local Similarity 57.1%; Pred. No. 6.6e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Y 1 KEYPTS 7
  1:1111
Db 275 VREFPTS 281

RESULT 14
US-09-782-874-2
Sequence 2: Application US/09/782874

Patent No. US20010023067A1
GENERAL INFORMATION:
APPLICANT: Wassenecker, Michael;
Riedel, Leonhard;
Schiebel, Manfred;
Sander, Heinz
TITLE OF INVENTION: NOVEL ACTA MOLECULES ENCODING
POLYPEPTIDES HAVING THE ENZYMATIC ACTIVITY OF AN
RNA-DIRECTED RNA POLYMERASE (RDRP)
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-09-782,874
FILING DATE: 08-Feb-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/411,583
FILING DATE: 05-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Haley, James F.
REGISTRATION NUMBER: 27,704
REFERENCE/DOCKET NUMBER: 000-1
TELECOMMUNICATION INFORMATION
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1114 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-782-874-2

Query Match: 77.8% Score 28; DB 10; Length 1114;
Best Local Similarity 57.1%; Pred. No. 9.9e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Y 1 KEYPTS 7
  1:1111
Db 843 VREFPTS 849

RESULT 15
US-09-144-838-14
Sequence 14: Application US/09/144838A
Patent No. US2002005196A1
GENERAL INFORMATION:
APPLICANT: Siani, Michael A.
APPLICANT: Wilken, Jill
APPLICANT: Simon, Reyna
APPLICANT: Kent, Stephen B.H.
TITLE OF INVENTION: Modular Protein Libraries and Methods of Preparation
FILE REFERENCE: GFPN-020/0105
CURRENT APPLICATION NUMBER: US/09/144,838A
CURRENT FILING DATE: 1998-08-31
EARLIER APPLICATION NUMBER: 60/071,620
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 14
LENGTH: 32
```

TYPE: PRI
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
-09-144-838-14

Query Match 75.0% Score 27; DB 10; Length 32;
Best Local Similarity: 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

1 IKEYF 5
|||||
23 IKEYF 27

arch completed: April 8, 2003, 12:06:21
b time : 13.8333 secs

GenCore version 5.1.4-P5-4578
Copyright (c) 1993 - 2003 Computer Ltd

M protein - protein search using SW model

April 8, 2003 11:56:19 Search time 12.0833 seconds
(without alignments)
17.163 Million cell updates/sec

Title: US-09-647-749A-1

Percent score: 25
Sequence: 1 EASY 5

Scoring table: BLOSUM62
Gapop 10.0 / Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40*

Fried. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	25	100.0	99 1	SY08_HUMAN
2	25	100.0	99 1	SY08_HUMAN
3	25	100.0	99 1	SY08_HUMAN
4	25	100.0	99 1	SY08_HUMAN
5	25	100.0	99 1	SY08_HUMAN
6	25	100.0	99 1	SY08_HUMAN
7	25	100.0	99 1	SY08_HUMAN
8	25	100.0	99 1	SY08_HUMAN
9	25	100.0	99 1	SY08_HUMAN
10	25	100.0	99 1	SY08_HUMAN
11	25	100.0	99 1	SY08_HUMAN
12	25	100.0	99 1	SY08_HUMAN
13	25	100.0	99 1	SY08_HUMAN
14	25	100.0	99 1	SY08_HUMAN
15	25	100.0	99 1	SY08_HUMAN
16	25	100.0	99 1	SY08_HUMAN
17	25	100.0	99 1	SY08_HUMAN
18	25	100.0	99 1	SY08_HUMAN
19	25	100.0	99 1	SY08_HUMAN
20	25	100.0	99 1	SY08_HUMAN
21	25	100.0	99 1	SY08_HUMAN
22	25	100.0	99 1	SY08_HUMAN
23	25	100.0	99 1	SY08_HUMAN
24	25	100.0	99 1	SY08_HUMAN
25	25	100.0	99 1	SY08_HUMAN
26	25	100.0	99 1	SY08_HUMAN
27	25	100.0	99 1	SY08_HUMAN
28	25	100.0	99 1	SY08_HUMAN
29	25	100.0	99 1	SY08_HUMAN
30	25	100.0	99 1	SY08_HUMAN
31	25	100.0	99 1	SY08_HUMAN
32	25	100.0	99 1	SY08_HUMAN
33	25	100.0	99 1	SY08_HUMAN

34	22	88.0	67	1	RGS3_RAI	P49797 ratnus mrv
35	22	88.0	99	1	SY08_HUMAN	009141 bos taurus
36	22	88.0	104	1	Y376_MOUSE	P47616 mycoplasma
37	22	88.0	142	1	YF45_STRIPY	P16848 streptococ
38	22	88.0	229	1	NEP1_CRYIA	Q664P2 canidia ala
39	22	88.0	237	1	Y4MC_PHEIN	P66662 rhizobium s
40	22	88.0	239	1	RNM_ASISA	P19791 aspergillus
41	22	88.0	244	1	IF6_SCHRO	Q94476 schizosarc
42	22	88.0	246	1	YCS0_BACIN	P42968 bacillus su
43	22	88.0	247	1	YE95_STRIPY	019767 schizosarc
44	22	88.0	256	1	TAM_PHEIN	Q94P73 rhizobium l
45	22	88.0	270	1	PP61_MOUSE	Q36655 mus musculu

ALIGNMENTS

RESULT 1

SY08_HUMAN STANDARD: PRT: 99 AA.

AC P80675: P78388: 01-DEC-1992 (Rel. 24, Created)

DT 01-NOV-1997 (Rel. 15, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Small inducible cytokine A8 precursor (CCL8) (Monocyte chemoattractant protein 2) (MCP-2) (Mx-2) (Chemokine (CCL8) protein 2) (CCL8)

GN SCYA8 OR SCYA10 OR MCP2

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Crania; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. AND VARIANT GIN 69.

RX MEDLINE:97237052; PubMed:9119400;

RA van Damme J., Fiten P., Nomiya H., Sakaki Y., Miura R., Yoshie O.,

RT "The human MCP-2 gene (SCYA8): cloning, sequence analysis, tissue expression, and assignment to the CC chemokine gene contig on

chromosome 17q11.2."

RL Genomics 40:323-331(1997).

RN [2]

RP SEQUENCE FROM N.A. AND VARIANT GIN 69.

RX MEDLINE:97237052; PubMed:9119400;

RA van Damme J., Fiten P., Nomiya H., Sakaki Y., Miura R., Yoshie O.,

RT "The human MCP-2 gene (SCYA8): cloning, sequence analysis, tissue expression, and assignment to the CC chemokine gene contig on

chromosome 17q11.2."

RL Genomics 40:323-331(1997).

RN [3]

RP SEQUENCE OF 23-69 FROM N.A.

RX MEDLINE:97237052; PubMed:9119400;

RA Chang H.C., Han P., Freeman G.J., Gillin J.D., Reinherz E.L.,

RT "Cloning and expression of a gamma interferon-inducible gene in

monocytes: a new member of a cytokine gene family."

PL Int. Immunol. 1:344-349(1989).

RN [4]

RP SEQUENCE OF 26-99.

RX MEDLINE:97237052; PubMed:9119400;

RA van Damme J., Fiten P., Nomiya H., Sakaki Y., Miura R., Yoshie O.,

RT "The human MCP-2 gene (SCYA8): cloning, sequence analysis, tissue expression, and assignment to the CC chemokine gene contig on

chromosome 17q11.2."

RL Genomics 40:323-331(1997).

RN [5]

RP SUBUNIT

RX MEDLINE:97053697; PubMed:8898111;

RA Kim Y.-S., Pajithanam P., Clardy J., Sykes B.D.,

RT "Structural characterization of a monomeric chemokine: monocyte

chemoattractant protein-3."

PL FERS Lett. 395:27-282(1996).

RL FUNCTION: CHEMOKINE FACTOR THAT ATTRACTS MONOCYTES, LYMPHOCYTES,

 C This SWISS-PROT entry is copyright. It is produced through a collaboration
 C between the Swiss Institute of Bioinformatics and the EMBL outstation -
 C the European Bioinformatics Institute. There are no restrictions on its
 C use by non-profit institutions as long as its content is in no way
 C modified and this statement is not removed. Usage by and for commercial
 C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 C or send an email to license@isb-sib.ch).
 C -----
 R EMBL: AEO01550; AAD06952.1; -
 R InterPro: IPR004552; AGP_AcylTrn.
 R InterPro: IPR002123; Acyltransferase.
 R Pfam: PF01553; Acyltransferase; 1.
 R TrEMBL: T16P00530; AGP_AcylTrn; 1
 W Phospholipid biosynthesis; Transferase; Acyltransferase;
 W Inner membrane; Complete proteome.
 Q SEQUENCE 237 AA; 27188 MW; E10F517D42A1731F CPG64;

 Y 1 LESYT 5
 Y 11111
 b 204 LESYT 208

 RESULT 4
 LSC_HELPY STANDARD; PRT; 240 AA.
 C 025603;
 T 15-DEC-1998 (Rel. 37, Created)
 T 15-DEC-1998 (Rel. 37, Last sequence update)
 T 15-JUN-2002 (Rel. 41, Last annotation update)
 E 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP
 E acyltransferase) (1-AGPAT) (Lysophosphatidic acid acyltransferase)
 E (IPAT)
 N PISC OR HP1348.
 N Helicobacter pylori (Campylobacter pylori).
 C Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 C Helicobacter.
 X NCBI_TaxID=2190;
 N [1]
 P SEQUENCE FROM N.A.
 P STRAIN=26695; ATCC: 43047;
 X MEDLINE=9394467; PubMed=9252185;
 A Tomb J.-F., White O., Kellavag A.R., Clayton R.A., Sutton G.G.,
 A Fleischmann P.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
 A Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
 A Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
 A McKenney K., Fitzgerald L.W., Lee N., Adams M.D., Hickey E.K.,
 A Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
 A Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Mallin E.,
 A Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 A Venter J.C.;
 T The complete genome sequence of the gastric pathogen Helicobacter
 T pylori.
 L Nature 388:539-547(1997).
 C ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.
 C -1- FUNCTION: CONVERTS LYSPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
 C ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.
 C -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -
 C CoA + 1,2-diacyl-sn-glycerol 3-phosphate
 C -1- PATHWAY: De novo phospholipid biosynthesis, second step.
 C -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
 C -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
 C ACYLTRANSFERASE FAMILY.
 C -----
 C This SWISS-PROT entry is copyright. It is produced through a collaboration
 C between the Swiss Institute of Bioinformatics and the EMBL outstation -
 C the European Bioinformatics Institute. There are no restrictions on its
 C use by non-profit institutions as long as its content is in no way
 C modified and this statement is not removed. Usage by and for commercial
 C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AEO0636; AAD08393.1; -
 DR TrEMBL: HP1348; -
 DR InterPro: IPR004552; AGP_AcylTrn.
 DR InterPro: IPR002123; Acyltransferase.
 DR Pfam: PF01553; Acyltransferase; 1.
 DR TrEMBL: T16R00530; AGP_AcylTrn; 1
 W Phospholipid biosynthesis; Transferase; Acyltransferase;
 W Inner membrane; Complete proteome.
 KW SEQUENCE 240 AA; 27745 MW; 22B1D0EB190BDD CPG64;

 QY 1 LESYT 5
 QY 11111
 DE 204 LESYT 208

 RESULT 5
 Y711_METUA STANDARD; PRT; 322 AA.
 AC 058121;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein M0711.
 GN M0711.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 CC Methanocaldococcaceae; Methanocaldococcus.
 NX NCBI_TaxID=2190;
 N [1]
 P SEQUENCE FROM N.A.
 P STRAIN=JAL-1 / DSM 2661 / ATCC 43047;
 RC MEDLINE=96337999; PubMed=8688087;
 RX Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kellavag A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.L.,
 RA Overbeek P., Kirkness E.F., Weinstock K.G., Merriam J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts J.M., Hurst M.A., Venter J.C., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT The complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii.
 PI Science 273:1058-1073(1996).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U67518; AAB98714.1; -
 CC DR TIGR: MJ0711; -
 CC KW Hypothetical protein; Transmembrane; Complete proteome.
 CC TRANSMEM 299 319 POTENTIAL.
 CC SEQUENCE 322 AA; 38631 MW; 07C1E417504F27 CPG64;

 QY 1 LESYT 5
 QY 11111
 DB 106 LESYT 110

 Query Match 100.0%; Score 25; DB 1; Length 322;
 Best Local Similarity 100.0%; Pred No. 43;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

of release factors".
 Mol. Biol. Evol. 17:982-989(2000).
 -1- FUNCTION: Directs the termination of nascent peptide synthesis (translation) in response to the termination codons UAA, UAG and UGA (by similarity).
 -1- SUBUNIT: HETERODIMER OF TWO SUBUNITS, ONE OF WHICH BINDS GTP.
 -1- SUBCELLULAR LOCATION: CYTOSOL; PLASMA MEMBRANE (BY SIMILARITY).
 -1- SIMILARITY: BELONGS TO THE EUKARYOTIC RELEASE FACTOR 1 FAMILY.

 This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch> or send an email to license@isb-sib.ch).

 EMBL: AF198107; AAF74402.1;
 HSSP: P46055; ITR9.
 InterPro: IPR005140; eRF1.
 InterPro: IPR005142; eRF1.
 InterPro: IPR005141; eRF1.
 Pfam: PF03463; eRF1_1; 1.
 Pfam: PF03464; eRF1_2; 1.
 Pfam: PF03465; eRF1_3; 1.
 TIGRFAMs: TIGR001001; eRF1; 1.
 PROSITE: PS00039; ZNF_C4; 1.
 PROTEIN BIOSYNTHESIS.
 SEQUENCE 457 AA: 51029 MW: 58185FFDIFQC943 CRC64:

 Query Match 100.0% Score 25; DB 1; Length 457;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 LESYT 5
 11111
 Db 30 LESYT 34

 RESULT 9
 RXRG_XENIA STANDARD; PRT: 470 AA.
 ID RXRG_XENIA STANDARD; PRT: 470 AA.
 AC P51129;
 DT 01-OCT-1996 (rel. 34, created)
 DT 01-OCT-1996 (rel. 34, last sequence update)
 DT 15-MAR-1999 (rel. 38, last annotation update)
 DE Retinoid acid receptor RXR-gamma.
 RN RXRG OR NR2B3.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae; Xenopodidae; Xenopus.
 NC NCBITaxID=8355;
 RN NCBITaxID=8355;
 RP SEQUENCE FROM N.A.
 BX MEDLINE:92356110; PubMed:312717;
 RA Blumberg P., Mangelsdorf D.J., Dyck J.A., Bittner P.A., Evans P.M., de Robertis E.M.;
 RT "Multiple retinoid-responsive receptors in a single cell: families of retinoid 'X' receptors and retinoid acid receptors in the Xenopus egg".
 RL Proc. Natl. Acad. Sci. U.S.A. 89:4341-4345(1992).
 CC -1- FUNCTION: INVOLVED IN RETINOIC ACID RESPONSE PATHWAY. BINDS 9-CIS RETINOIC ACID (9CRA) (BY SIMILARITY).
 CC -1- SUBUNIT: HOMODIMER OR HETERODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- DEVELOPMENTAL STAGE: IT IS SYNTHESIZED DURING OOGENESIS AND PERSISTS IN THE CLEAVING EMBRYO AT APPROXIMATELY CONSTANT LEVELS UNTIL IT IS DEGRADED JUST BEFORE GASTRULATION.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS, A MODULATING N-TERMINAL DOMAIN, A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NR2 SUBFAMILY.

 This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch> or send an email to license@isb-sib.ch).

 EMBL: U1143; NOT_ANNOTATED_CDS
 HSSP: P19793; ZNF.
 TRANSFAC: T01960;
 InterPro: IPR000536; Hormone_rec_1; 1.
 InterPro: IPR001723; Steroid_receptor; 1.
 InterPro: IPR001729; Znf_C4steroid
 Pfam: PF00104; hormone_rec; 1.
 Pfam: PF00105; Zf-C4; 1.
 PRINTS: PR00398; STEROIDOMER.
 PRINTS: PR00047; STEROIDOMER.
 PRODOM: PD000935; Znf_C4steroid; 1
 SMART: SM00430; HOL1; 1.
 SMART: SM00399; ZNF_C4; 1.
 PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 DR PROSITE: Transcription regulation: DNA-binding; Nuclear protein; Zinc-finger; Multigene family.
 KW ZINC-FINGER; MULTIGENE FAMILY.
 FT DOMAIN 1 145 MODULATING (BY SIMILARITY).
 FT DNA_BIND 146 211 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 146 166 C4-TYRE.
 FT ZN_FING 182 205 C4-TYRE.
 FT DOMAIN 212 275 HINGE.
 FT DOMAIN 276 470 LIGAND BINDING (BY SIMILARITY).
 FT PT
 SEQUENCE 470 AA: 51925 MW: 103172053CF9A5F CRC64:

 Query Match 100.0% Score 25; DB 1; Length 470;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 LESYT 5
 11111
 Db 408 LESYT 412

 RESULT 10
 MURD_TREPA STANDARD; PRT: 532 AA
 ID MURD_TREPA STANDARD; PRT: 532 AA
 AC O83873;
 DT 15-DEC-1998 (rel. 37, created)
 DT 15-DEC-1998 (rel. 37, last sequence update)
 DT 15-JUN-2002 (rel. 41, last annotation update)
 DE UDP-N-acetylaminotyrosylalanine-D-glutamate ligase (EC 6.3.2.9) (UDP-N-acetylaminotyrosyl-L-alanyl-D-glutamate synthetase) (D-glutamic acid adding enzyme).
 RN MURD OR TP0903
 OS Treponema pallidum.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
 NC NCBITaxID=160;
 RN NCBITaxID=160;
 RP SEQUENCE FROM N.A.
 BX MEDLINE:9672770; PubMed:9665876;
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G., O-Jasbi R., Gwin M., Hickey E.K., Lyden P., Potluri K.A., Solberg E., Hardham T.M., Melton M.P., Salzberg S., Peterson J., Khalak H., Richardson D., Howell I.K., Chidambaram M., Osterback T., McDonald L., Attlich P., Pownall G., Cotton M.D., Fujii C., Gailand S., Hitch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith B.D., Venter J.C.;
 RT "Complete genome sequence of Treponema pallidum, the syphilis spirochete".
 RL Science 291:375-382(1998).
 CC -1- FUNCTION: CELL WALL FORMATION. CATALYZES THE ADDITION OF D-GLUTAMATE TO THE NUCLEOTIDE PRECURSOR UDP-N-ACETYLAMINOYL-L-ALANINE (UAA) (BY SIMILARITY).

1- CATALYTIC ACTIVITY: ATP + UDP-N-acetylglutamate-L-alanine + D-glutamate - ADP + phosphate + UDP-N-acetylglutamate-L-alanyl-D-alaninate

1- PATHWAY: Peptidoglycan biosynthesis

1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity)

1- SIMILARITY: BELONGS TO THE MURDEF FAMILY

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch)

EMBL: AF001259; AAC65856.1; -

HSSP: P14900; IEH.

TIGR: TP0903; -

InterPro: IPR000713; Mur_Ligase.

InterPro: IPR004101; Mur_Ligase_C.

Pfam: PF01225; Mur_Ligase_2.

Pfam: PF02875; Mur_Ligase_C_1.

TIGRPFAM: TIGR01087; MurL_1.

Peptidoglycan synthetase; Cell wall; Cell division; Ligase

ATP-binding; Complete proteome.

NP_BIND 124 130

SEQUENCE 532 AA: 57447 MW: 970350; AAF476A6; C6064

Query Match 100.0% Score 25; DB 1; Length 532

Best Local Similarity 100.0% Pred. No. 76

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0

1 LESYT 5

|||||

385 LESYT 389

RESULT 11

570_CHICK STANDARD: PRT: 634 AA.

1- AUG-1988 (Rel. 08, Created)

1- AUG-1988 (Rel. 08, Last sequence update)

1- FEB-1994 (Rel. 28, Last annotation update)

Heat shock 70 kDa protein (HSP70).

Gallus gallus (Chicken)

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianidae; Gallus.

NCBI_TaxID: 9031.

SEQUENCE FROM N.A.

MEDLINE: 86304452; PubMed: 3017985;

Morimoto R.I., Hunt C., Huang S.Y., Berg K.L., Banerji S.S.;

"Organization, nucleotide sequence, and transcription of the chicken HSP70 gene."

J. Biol. Chem. 261:12692-12699(1986).

1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch)

EMBL: J02579; AAA48825.1; -

PIR: A25646; A25646.

HSSP: P08109; ICKR

InterPro: IPR001023; HSP70.

Pfam: PF00012; HSP70_1

PRINTS: PR00301; HEATSHOCK70.

DR PROSITE: PS00089; HSP70_1.

DR PROSITE: PS00297; HSP70_1.

DR PROSITE: PS00329; HSP70_2; 1.

DR PROSITE: PS01036; HSP70_3; 1.

KW ATP-binding; Heat shock; Multigene family.

SO SEQUENCE 634 AA: 69750 MW: 427077; P080D3; A6AEB; C6064;

Query Match 100.0% Score 25; DB 1; Length 634

Best Local Similarity 100.0% Pred. No. 93

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0

1 LESYT 5

|||||

545 LESYT 549

RESULT 12

HS72_HUMAN STANDARD: PRT: 639 AA.

1- OCT-1996 (Rel. 34, Created)

01-OCT-1996 (Rel. 34, Last sequence update)

15-OCT-2002 (Rel. 42, Last annotation update)

Heat shock-related 70 kDa protein 2 (Heat shock 70 kDa protein 2); HSPA2.

OS Homo sapiens (Human);

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NCBI_TaxID: 9606;

1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY

SEQUENCE FROM N.A.

MEDLINE: 96130116; PubMed: 7829106;

Bongycastle L.L.C., Yu C.-E., Hunt C.F., Trask B.J., Clancy K.P., Weber J.L., Patterson D., Schellenberg G.D.;

"Cloning, sequencing, and mapping of the human chromosome 14 heat shock protein gene (HSPA2)."

Genomics 23:85-93(1994).

1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY

SEQUENCE FROM N.A.

Goraliski T.J., Krensky A.M.;

Submitted (Apr-1996) to the EMBL/Genbank/CCDB databases

1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY

PREEXISTENT PROTEINS AGAINST AGGREGATION AND MEDATE THE FOLDING OF NEWLY TRANSLATED POLYPEPTIDES IN THE CYTOSOL AS WELL AS WITHIN ORGANELLERS. THESE CHAPERONES PARTICIPATE IN ALL THESE PROCESSES THROUGH THEIR ABILITY TO RECOGNIZE NONNATIVE CONFORMATIONS OF OTHER PROTEINS. THEY BIND EXTENDED PEPTIDE SEGMENTS WITH A NET HYDROPHOBIC CHARACTER EXPOSED BY POLYPEPTIDES DURING TRANSLATION AND MEMBRANE TRANSLOCATION, OR FOLLOWING STRESS-INDUCED DAMAGE.

1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch)

EMBL: L26356; AAA52698.1; -

EMBL: U56725; AAD11466.1; -

EMBL: U01439; AAC50076.1; -

HSSP: P19120; 3HSC

Genew: HSCW: 5235; HSPA2.

MIM: 140560;

InterPro: IPR001023; HSP70.

Pfam: PF00012; HSP70_1.

InterPro: IPR005259; Pfam: PF00271; Helicase_C: 1.
 SMART: SM00487; DEXDc: 1.
 TIGRfams: TIGR00595; Pfam: 1.
 DNA replication: DNA-binding; Helicase; Primosome;
 Zinc-finger; Complete proteome.
 NP_BIND 241 248 ATP (POTENTIAL).
 SITE 337 340 DEH_BOX.
 ZN_FING 458 470 C4-TYPE (POTENTIAL).
 ZN_FING 485 501 C4-TYPE (POTENTIAL).
 SEQUENCE 753 AA: 84764 MW: 9245277.6 CAAB05.CRC64;
 Query Match 100.0%; Score 25; DB 1; Length 753;
 Best Local Similarity 100.0%; Pred. No. 1; 1e-02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LESYT 5
 |||||
 379 LESYT 383

SULT 15
 IA_CHLTR STANDARD: PRI: 753 AA.
 084783;
 16-OCT-2001 (Rel. 40, Created)
 16-OCT-2001 (Rel. 40, Last sequence update)
 16-OCT-2001 (Rel. 40, Last annotation update)
 Primosomal protein N (Replication factor Y).
 PRIA OR C1778.
 Chlamydia trachomatis.
 Bacteria: Chlamydiales; Chlamydiaceae; Chlamydia.
 NCBI_TaxID=813;
 (1)

SEQUENCE FROM N.A.
 STRAIN-D/OW-3/CX: PubMed:9784186;
 MEDLINE=99000809; Stephens R.S., Kaiman S., Lammell C.J., Fan J., Marathe R., Aravind L., Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V., Davis R.W.,
 Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis.";
 Science 282:754-759(1998).

-1- FUNCTION: RECOGNIZES A SPECIFIC HAIRPIN SEQUENCE ON PRIA ssDNA
 THIS STRUCTURE IS THEN RECOGNIZED AND BOUND BY PROTEINS PRIA AND
 ACTIONS OF DNAB, DNAC, DNAT AND PRIMASE. PRIA THEN FUNCTIONS AS A
 HELICASE WITHIN THE PRIMOSOME (BY SIMILARITY).
 -1- SIMILARITY: BELONGS TO THE HELICASE FAMILY. PRIA SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/Edlib/Edlib.html> or send an email to license@ebi.ac.uk).

EMBL: AE001350; AAC6873.1;
 InterPro: IPR001410; DEAD.
 InterPro: IPR001650; Helicase_C.
 InterPro: IPR005259; PRIA.
 Pfam: PF00271; Helicase_C: 1.
 SMART: SM00487; DEXDc: 1.
 SMART: SM00490; HELICc: 1.
 TIGRfams: TIGR00595; Pfam: 1.
 DNA replication: DNA-binding; ATP-binding; Helicase; Primosome;
 Zinc-finger; Complete proteome.
 NP_BIND 241 248 ATP (POTENTIAL).
 SITE 337 340 DEH_BOX.
 ZN_FING 458 470 C4-TYPE (POTENTIAL).
 ZN_FING 485 501 C4-TYPE (POTENTIAL).
 SEQUENCE 753 AA: 84831 MW: 9216154.38 CAAB05.CRC64;

Query Match 100.0%; Score 25; DB 1; Length 753;
 Best Local Similarity 100.0%; Pred. No. 1; 1e-02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LESYT 5
 |||||
 DB 379 LESYT 383
 Search Completed: April 8, 2003, 12:02:14
 Job time: 15.0833 secs

New peptides useful for inhibiting human immuno-deficiency virus type 1 (HIV-1) gp120 induced neuronal cell death

Claim 1: Page 11; 16pp: English.

AAV49695 and AAV49696 represent peptides which inhibit HIV-1 gp120 induced neuronal cell death. Pharmaceutical compositions containing the peptides are useful for treating symptoms caused by neuronal cell loss. Such conditions especially associated with HIV infection include encephalopathies, neuropathies, memory loss, dementia, depression, psychosis and opportunistic infections. The peptides act as antagonists of gp120-mediated neurotoxicity and subsequent neuronal degeneration. This enables therapeutic treatment of HIV infection and other inflammatory neurological diseases, including multiple sclerosis, tropical spastic paraparesis and Alzheimer's disease.

Sequence 7 AA:

Query Match 100.0% Score 46; Db 20; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7; Ho-05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 IKEYPTS 7
 |||||
 1 IKEYPTS 7

24-JUL-2002 (first entry)
 Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3833.

ABP38988:

24-JUL-2002 (first entry)
 Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3833.
 Staphylococcus epidermidis: open reading frame: ORF, bacterial infection, antibacterial; gene therapy.

Staphylococcus epidermidis:

US6380370-B1.

30-APR-2002.

13-AUG-1998, 9805 0134001.

14-AUG-1997, 9705-055779P.

08-NOV-1997, 9705-064964P.

(GENO-) GENOME THERAPEUTICS CORP.

Doucette-Stamm LA, Bush D;

MP1: 2002-38125/41.

N-PSDB: ABN81533.

Novel isolated nucleic acid encoding a Staphylococcus epidermidis polypeptide, useful for diagnosing and treating bacterial infections.

Disclosure: SEQ ID 3833; 26pp: English.

ABN90538 to ABN93374 represent staphylococcus epidermidis open reading frame (ORF) nucleic acid sequences which encode the amino acid sequences given in ABN95124 to ABN97960. The S. epidermidis sequences have antibacterial activity and can be used in gene therapy. The sequences can also be used in the diagnosis and treatment of bacterial infections, particularly S. epidermidis infections. The sequences can be used to screen for compounds able to interfere with the S. epidermidis life cycle or inhibit S. epidermidis infection.

N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the

CC USPRO web site.
 XX Sequence 204 AA;

Query Match 88.9%; Score 46; Db 23; Length 204;
 Best Local Similarity 71.4%; Pred. No. 97;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKEYPTS 7
 |||||
 Db 195 VKFFFA 201

RESULT 3
 AAW3792
 ID AAW3792 standard; Protein; 422 AA.

XX AAW3792;

XX 10-NOV-1997 (first entry)

XX Cytidine monophospho-N-acetylneuraminic acid hydroxylase.

XX CMP-Neu5Ac hydroxylase; CMP-Neu5Ac hydroxylase; enzyme activity;

XX glycoprotein; cytochrome b5; tumour human; pig.

XX Sus scrofa.

XX W05703200-A1.

XX 30-JAN-1997.

XX 05-JUL-1996; 96WO-EP02966.

XX 22-SEP-1995; 95EP-0114930.

XX 07-JUL-1995; 95JP-0110609.

XX (POPF) POPFSTRCEP MANNHFM GMBH.

XX Haselbeck A, Honold K, Kelm S, Schaner R, Schlenzka W;

XX Shaw L;

XX MP1: 1997-119055/11.

XX N-PSDB: AAT78596.

XX Claim 2: Page 32-33; 44pp: English.

XX The present sequence represents a now isolated protein cytidine

XX monophospho-N-acetylneuraminic acid hydroxylase, which is uncontaminated

XX by cytochrome b5 or cytochrome b5 reductase. The protein is produced by

XX exogenous DNA in prokaryotic or eukaryotic cells and it catalyses

XX conversion of CMP Neu5Ac to CMP Neu5Gc (Neu5Ac and Neu5Gc - N-acetyl or

XX glycosyl-neuraminic acid respectively). Oligonucleotides which hybridise

XX to cDNA encoding the protein are used to detect such nucleic acids by

XX hybridisation assays, and are used to inhibit expression of the protein,

XX i.e. as antisense molecules for treatment of tumours, particularly those

XX of high metastatic potential. Cells that have been modified so that they

XX no longer produce the protein are used to produce Neu5Gc-free

XX glycometabolites for therapeutic use. The DNA encoding the protein is

XX dominant, can be used to diagnose tumours. Glycoproteins that lack Neu5Gc

XX more closely resemble human products and are suitable for long term

XX therapy since they contain no non-human epitopes (Neu5Gc containing

XX proteins are immunogenic in humans).

Query Match 88.9%; Score 46; Db 18; Length 422;

Best Local Similarity: 100.0%; Pred. No. 1.9e+02;
Matches: 6; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

1 IKEYFT 6
|||||
303 IKEYFT 308

RESULT 4

AAW10735 standard; Protein: 422 AA.

AAW10735;

30-SEP-1997 (first entry)

Porcine CMP-N-acetyl-neuraminic acid hydroxylase.

N-acetyl-neuraminic acid; Neu5Ac; sialic acid; porcine; CMP; cytidine monophosphate; antisense therapy; recombinant production; N-glycolylneuraminic acid side chain deficient; erythropoietin.

Sus scrofa.

EP752474-A1.

08-JAN-1997.

07-JUL-1995; 95EP-0110609.

07-JUL-1995; 95EP-0110609.

(BOEF) BOEHRINGER MANNHEIM GMBH.

Haselebeck A, Honold K, Kelm S, Schauer R, Schlenszka W;

Shaw L;

WPI: 1997-067451/07.

N-PSDB: AAI48588.

Porcine CMP-N-acetyl-neuraminic acid hydroxylase - catalyses conversion of CMP-Neu5Ac to CMP-Neu5Gc; antisense sequences of DNA encoding CMP-N-acetyl-neuraminic acid hydroxylase are used for prodn. of modified glycoproteins

Claim 2: Page 14-19; 24pp; English.

AAW10735 represents a porcine cytidine mono phosphate (CMP; N-acetyl-neuraminic acid (CMP-Neu5Ac) hydroxylase enzyme. The enzyme catalyses the conversion of CMP-Neu5Ac to CMP-Neu5Gc (cytidine monophosphate-N-glycolylneuraminic acid). Antisense sequences of this cDNA sequence are used for the production of CMP-Neu5Gc hydroxylase-deficient mammalian cells. The modified cells are useful for the production of recombinant glycoproteins lacking N-glycolylneuraminic acid side chains, e.g. erythropoietin.

Sequence 422 AA;

Query Match: 88.9%; Score 32; DB 18; Length 422;

Best Local Similarity: 100.0%; Pred. No. 1.9e+02;

Matches: 6; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

1 IKEYFT 6
|||||

303 IKEYFT 308

RESULT 5

ABP26246 standard; Protein: 728 AA.

ABP26246;

ABP26246;

ABP26246;

ABP26246;

ABP26246;

ABP26246;

02-JUL-2002 (first entry)

Streptococcus polypeptide SEQ ID NO 1668.

Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae; group A streptococcus; Streptococcus pyogenes; antibacterial; antiinflammatory; infection; vaccine; meningitis; gene therapy.

Streptococcus pyogenes

W0200234771-A2.

02-MAY-2002.

24-OCT-2000; 2001W-060479.

27-OCT-2000; 2000GR-0026333.

24-NOV-2000; 2000GB-0028727.

07-MAR-2001; 2001SH-0005640.

(CHIR-) CHIRON SPA.

(GENO-) INST GENOMIC RES.

Telford J, Maignani V, Margarit Ros YL, Grandi G, Fraser C;

Tectell H;

WPI: 2002-352536/38.

N-PSDB: ABN66877.

Claim 1: Page 3318; 4525pp; English.

The invention relates to a protein (ABP26246) from group B streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in the specification. The proteins have antibacterial and antiinflammatory activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and antibodies that bind (1) are used in the manufacture of medicaments for the treatment or prevention of infection of disease caused by Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

Nucleic acids encoding (1) are used to detect Streptococcus in a biological sample. (1) is used to determine whether a compound binds to (1). A composition comprising (1) or a nucleic acid encoding (1), may be used as a vaccine or diagnostic composition. The disease caused by Streptococcus that is prevented or treated may be meningitis, Nucleic acid encoding (1) may be used to recombinantly produce (1) and may be used in gene therapy. Antibodies to (1) are used for affinity chromatography, immunoassays, and distinguishing/identifying Streptococcus proteins.

Sequence 728 AA;

Query Match: 88.9%; Score 32; DB 23; Length 728;

Best Local Similarity: 71.4%; Pred. No. 3.3e+02;

Matches: 5; Conservative: 2; Mismatches: 0; Indels: 0; Gaps: 0;

1 IKEYFTS 7
|||||

131 VKDYFTS 137

RESULT 6

AAAG70745 standard; Protein: 1460 AA.

AAAG70745;

AAAG70745;

AAAG70745;

AAAG70745;

AAAG70745;

AAAG70745;

AAAG70745;

AAAG70745;

AAAG70745;

AAAG70745;

AAAG70745;

AAAG70745;

Yeast; fungus; apoptosis; infection; proliferative disease;
vaccine; autoimmune disease; ischaemia; neurodegeneration;

Saccharomyces cerevisiae.

WO200102550-A2.

11-JAN-2001.

03-JUL-2000; J06060-HE-00077.

01-JUL-1999; 99EP-0870141.

(JANC) JANSSEN PHARM NV.

Contreras RH, De Backer MD, Luyten WHML, Malcorps IKL,
Nellissen BM, Reekmans RJ;

WPI: 2001-367442/96.

N-PSDB: AAH29781.

Yeast and fungal nucleic acids encoding proteins involved in a pathway
leading to programmed cell death, useful for treating proliferative
disorders, yeast and fungal infections, or for preventing apoptosis in
certain diseases.

Claim 1: Fig 1: 218pp: English.

The present invention provides the protein and coding sequences of a
number of apoptosis associated proteins from the yeast *Saccharomyces*
cerevisiae and the fungus *Candida albicans*. These can be used to identify
treatments for fungal and yeast infections, for proliferative diseases
and for apoptosis related diseases such as autoimmune diseases, ischaemia
and neurodegeneration. The present sequence is one of the *S. cerevisiae*
proteins of the invention.

Sequence 1460 AA:

Query Match 86.1% Score 31; DB 22; Length 1460;
Best Local Similarity 71.4% Pred. No. 1e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

1 KEYPTS 7

1111111

1190 LKEYPTS 1196

SULT 7

129339

AA029339 standard; Protein: 150 AA.

AA029339;

18-DEC-2001 (first entry)

Novel mar regulated protein (NIMR) #11.

mar regulated polypeptide; NIMR; microbial infection; antibacterial.

Escherichia coli.

WO200170776-A2.

27-SEP-2001.

08-MAR-2001; 2001WO-US07478.

10-MAR-2000; 2000US-188362P.

(TUFT) TUFTS COLLEGE.

Levy SB, Balhosa TM, Alekshun MN;

XX WPI: 2001-60376/49.
XX N-PSDB: AAS46240.

XX Identifying compounds that modulate a newly identified mar regulated

PT polypeptide activity, useful as antimicrobial compounds. Involves

PS contacting the polypeptide with a test compound.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Sequence 150 AA:

Query Match 83.3% Score 31; DB 22; Length 150;
Best Local Similarity 83.3% Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEYFT 6

DB 137 LKEYFT 142

RESULT 8

AA098376

ID AA098376 standard; Protein: 150 AA.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Sequence 150 AA:

Query Match 83.3% Score 31; DB 22; Length 150;
Best Local Similarity 83.3% Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEYFT 6

DB 137 LKEYFT 142

RESULT 8

AA098376

ID AA098376 standard; Protein: 150 AA.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Sequence 150 AA:

Query Match 83.3% Score 31; DB 22; Length 150;
Best Local Similarity 83.3% Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEYFT 6

DB 137 LKEYFT 142

RESULT 8

AA098376

ID AA098376 standard; Protein: 150 AA.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Sequence 150 AA:

Query Match 83.3% Score 31; DB 22; Length 150;
Best Local Similarity 83.3% Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEYFT 6

DB 137 LKEYFT 142

RESULT 8

AA098376

ID AA098376 standard; Protein: 150 AA.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Sequence 150 AA:

Query Match 83.3% Score 31; DB 22; Length 150;
Best Local Similarity 83.3% Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEYFT 6

DB 137 LKEYFT 142

RESULT 8

AA098376

ID AA098376 standard; Protein: 150 AA.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Sequence 150 AA:

Query Match 83.3% Score 31; DB 22; Length 150;
Best Local Similarity 83.3% Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEYFT 6

DB 137 LKEYFT 142

RESULT 8

AA098376

ID AA098376 standard; Protein: 150 AA.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Sequence 150 AA:

Query Match 83.3% Score 31; DB 22; Length 150;
Best Local Similarity 83.3% Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEYFT 6

DB 137 LKEYFT 142

RESULT 8

AA098376

ID AA098376 standard; Protein: 150 AA.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Sequence 150 AA:

Query Match 83.3% Score 31; DB 22; Length 150;
Best Local Similarity 83.3% Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEYFT 6

DB 137 LKEYFT 142

RESULT 8

AA098376

ID AA098376 standard; Protein: 150 AA.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

(1) have antibacterial and antibiotic activities, and can be used in gene therapy. Expression of (1) in a microorganism inhibits proliferation of the microorganism, and the manufactured antibiotic is useful for reducing the activity or level of a gene product required for proliferation of a microorganism in a subject, specifically humans. The nucleic acids that inhibit bacterial growth or proliferation can be used as antisense therapeutics for killing bacteria. In addition to therapeutic applications, the nucleic acid sequences complementary to sequences required for proliferation can be used as diagnostic tools. For example, nucleic acid probes complementary to proliferation-regulated sequences that are specific for particular species of microorganisms can be used as probes to identify particular microorganism species in clinical specimens. AAH81246 to AAH81487 encode the *Escherichia coli* proteins given in AAG94219 to AAG94431, and AAH81488 to AAH81491 represent oligonucleotides, which are used in the exemplification of the present invention.

Sequence 150 AA:

Query Match 81.3% Score 30: DB 22: Length 150:

Best Local Similarity 81.3% Frd- No. 1 8e-02: Mismatches 0: Gaps 0:

Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

1 IKERYT 6
137 IKERYT 142

ESUT 9
AAG42212
AAG42212 standard: Protein: 258 AA

AAG42212:

18-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO. 52616.

Protein identification: signal transduction pathway: metabolic pathway:
hybridisation assay: genetic mapping: gene expression control: promoter:
termination sequence.

Arabidopsis thaliana.

EP103405-A2.

06-SEP-2000.

25-FEB-2000: 2000EP 0303439.

25-FEB-1999: 99US-011825.
05-MAR-1999: 99US-0123180.
09-MAR-1999: 99US-0123548.
23-MAR-1999: 99US-0125788.
25-MAR-1999: 99US-0126264.
29-MAR-1999: 99US-0126785.
01-APR-1999: 99US-0127462.
06-APR-1999: 99US-0128234.
08-APR-1999: 99US-0128714.
16-APR-1999: 99US-0129845.
19-APR-1999: 99US-0130077.
21-APR-1999: 99US-0130449.
23-APR-1999: 99US-0130510.
28-APR-1999: 99US-0130891.
30-APR-1999: 99US-0131449.
30-APR-1999: 99US-0132048.
04-MAY-1999: 99US-0132407.
05-MAY-1999: 99US-0132484.
06-MAY-1999: 99US-0132485.
07-MAY-1999: 99US-0132486.
11-MAY-1999: 99US-0132487.
99US-0132488.
99US-0132489.
99US-0132490.
99US-0132491.
99US-0132492.
99US-0132493.
99US-0132494.
99US-0132495.
99US-0132496.
99US-0132497.
99US-0132498.
99US-0132499.
99US-0132500.
99US-0132501.
99US-0132502.
99US-0132503.
99US-0132504.
99US-0132505.
99US-0132506.
99US-0132507.
99US-0132508.
99US-0132509.
99US-0132510.
99US-0132511.
99US-0132512.
99US-0132513.
99US-0132514.
99US-0132515.
99US-0132516.
99US-0132517.
99US-0132518.
99US-0132519.
99US-0132520.
99US-0132521.
99US-0132522.
99US-0132523.
99US-0132524.
99US-0132525.
99US-0132526.
99US-0132527.
99US-0132528.
99US-0132529.
99US-0132530.
99US-0132531.
99US-0132532.
99US-0132533.
99US-0132534.
99US-0132535.
99US-0132536.
99US-0132537.
99US-0132538.
99US-0132539.
99US-0132540.
99US-0132541.
99US-0132542.
99US-0132543.
99US-0132544.
99US-0132545.
99US-0132546.
99US-0132547.
99US-0132548.
99US-0132549.
99US-0132550.
99US-0132551.
99US-0132552.
99US-0132553.
99US-0132554.
99US-0132555.
99US-0132556.
99US-0132557.
99US-0132558.
99US-0132559.
99US-0132560.
99US-0132561.
99US-0132562.
99US-0132563.
99US-0132564.
99US-0132565.
99US-0132566.
99US-0132567.
99US-0132568.
99US-0132569.
99US-0132570.
99US-0132571.
99US-0132572.
99US-0132573.
99US-0132574.
99US-0132575.
99US-0132576.
99US-0132577.
99US-0132578.
99US-0132579.
99US-0132580.
99US-0132581.
99US-0132582.
99US-0132583.
99US-0132584.
99US-0132585.
99US-0132586.
99US-0132587.
99US-0132588.
99US-0132589.
99US-0132590.
99US-0132591.
99US-0132592.
99US-0132593.
99US-0132594.
99US-0132595.
99US-0132596.
99US-0132597.
99US-0132598.
99US-0132599.
99US-0132600.

21-JUL-1999; 9905-0145224
 26-JUL-1999; 9905-0145226
 27-JUL-1999; 9905-0145231
 27-JUL-1999; 9905-0145238
 27-JUL-1999; 9905-0145241
 28-JUL-1999; 9905-0145251
 02-AUG-1999; 9905-0145386
 02-AUG-1999; 9905-0145388
 02-AUG-1999; 9905-0145389
 03-AUG-1999; 9905-0145388
 04-AUG-1999; 9905-0145388
 04-AUG-1999; 9905-0145388
 05-AUG-1999; 9905-0145388
 05-AUG-1999; 9905-0145388
 06-AUG-1999; 9905-0145388
 06-AUG-1999; 9905-0145388
 09-AUG-1999; 9905-0145388
 09-AUG-1999; 9905-0145388
 10-AUG-1999; 9905-0145388
 11-AUG-1999; 9905-0145388
 12-AUG-1999; 9905-0145388
 13-AUG-1999; 9905-0145388
 13-AUG-1999; 9905-0145388
 17-AUG-1999; 9905-0145388
 18-AUG-1999; 9905-0145388
 20-AUG-1999; 9905-0145388
 20-AUG-1999; 9905-0145388
 20-AUG-1999; 9905-0145388
 23-AUG-1999; 9905-0145388
 23-AUG-1999; 9905-0145388
 25-AUG-1999; 9905-0150566
 26-AUG-1999; 9905-0150884
 27-AUG-1999; 9905-0151065
 27-AUG-1999; 9905-0151066
 27-AUG-1999; 9905-0151080
 30-AUG-1999; 9905-0151303
 31-AUG-1999; 9905-0151438
 01-SEP-1999; 9905-0151930
 07-SEP-1999; 9905-0152363
 10-SEP-1999; 9905-0153070
 13-SEP-1999; 9905-0153758
 15-SEP-1999; 9905-0154018
 16-SEP-1999; 9905-0154039
 20-SEP-1999; 9905-0154779
 22-SEP-1999; 9905-0155139
 23-SEP-1999; 9905-0155486
 24-SEP-1999; 9905-0155659
 28-SEP-1999; 9905-0156458
 29-SEP-1999; 9905-0156596
 04-OCT-1999; 9905-0157117
 05-OCT-1999; 9905-0157753
 06-OCT-1999; 9905-0157865
 07-OCT-1999; 9905-0158029
 08-OCT-1999; 9905-0158442
 12-OCT-1999; 9905-0158369
 13-OCT-1999; 9905-0158369
 13-OCT-1999; 9905-0158369
 13-OCT-1999; 9905-0158369
 14-OCT-1999; 9905-0159295
 14-OCT-1999; 9905-0159329
 14-OCT-1999; 9905-0159330
 14-OCT-1999; 9905-0159331
 14-OCT-1999; 9905-0159337
 14-OCT-1999; 9905-0159338
 18-OCT-1999; 9905-0159584
 21-OCT-1999; 9905-0160741
 21-OCT-1999; 9905-0160767
 21-OCT-1999; 9905-0160768
 21-OCT-1999; 9905-0160770
 21-OCT-1999; 9905-0160814
 21-OCT-1999; 9905-0160815
 22-OCT-1999; 9905-0160980
 22-OCT-1999; 9905-0160981

PP 22-OCT-1999; 9905-0160989
 PP 25-OCT-1999; 9905-0161404
 PP 25-OCT-1999; 9905-0161405
 PP 25-OCT-1999; 9905-0161406
 PP 26-OCT-1999; 9905-0161359
 PP 26-OCT-1999; 9905-0161360
 PP 26-OCT-1999; 9905-0161361
 PP 28-OCT-1999; 9905-0161920
 PP 28-OCT-1999; 9905-0161992
 PP 28-OCT-1999; 9905-0161993
 PP 29-OCT-1999; 9905-0162142

 Query Match 83.3% Score 108.21; Length 258;
 Best Local Similarity 71.4% Pred. No. 36.02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

 QY 1 KEYPTS 7
 DB 22 KEYPTS 28

 RESULT 10
 ID ABB53504 Standard: Protein; 402 AA.
 XX ABB53504;
 AC
 XX
 DT 16-MAY-2002 (first entry)
 XX
 XX Lactococcus lactis protein rpa
 PF
 XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 OS
 XX Lactococcus lactis IL1403.
 PN FR2807446-A1.
 XX
 PD 12-OCT-2001.
 XX
 PF 11-APP-2000; 2000FP-0004630.
 XX
 PF 11-APP-2000; 2000FP-0004630.
 XX
 PA (INPG) INPA INST NAT RECH AGRONOMIQUE.
 XX
 PI Biotine A, Sorokine A, Renault F, Ehrlich SD;
 XX
 DR WPI: 2002-043418/06.
 PT New nucleotide sequence useful in the identification of Lactococcus
 PT lactis and related species -
 XX
 PS Claim 6: SEQ ID No 206; 2504FP; French.
 XX
 CC The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
 CC nucleic acid sequence is useful in the detection and/or amplification of
 CC nucleic acid sequence, particularly to identify Lactococcus lactis or
 CC related species. The products of the invention are useful for the
 CC biosynthesis or biodegradation of a composition of interest. The
 CC invention helps research in lactic bacteria, particularly useful in the
 CC production of yogurt and cheese.
 CC Note: The sequence data for this patent is based on equivalent patent
 CC WO2001/7334 (published 18-OCT-2001) which is available in electronic
 CC format directly from WIPO at http://www.int/pat/published/Pct_sequences.
 CC
 XX Sequence 402 AA:

 Query Match 83.3% Score 108.21; Length 402;
 Best Local Similarity 85.7% Pred. No. 4.6e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 QY 1 KEYPTS 7

|||||
b 169 IKEYLTS 175

RESULT 11

BB92997
D ABB92997 standard; Protein: 573 AA.

C ABB92997;

X 31-MAY-2002 (first entry)

X Herbicidally active polypeptide SEQ ID NO 2208.

X Herbicidal; plant: agriculture; herbicide.

X Arabidopsis thaliana.

X WQ200210210-A2.

X 07-FEB-2002.

X 29-AUG-2001; 2001WO-EP09892

X 28-AUG-2001; 2001WO-EP09892.

X (FARB) BAYER AG.

X Tietjen K, Weidler M;

X WPI: 2002-269010/31.

X Identifying plant target proteins for herbicidally active compounds,

X comprising aligning and comparing nucleic acid or amino acid sequences

X from plant with nucleic acid or amino acid sequences from non-plant

X organisms.

X Claim 5: SEQ ID NO 2208; 261pp + Sequence Listing; English.

X The invention relates to identifying target proteins

X (ABB90790-ABB94016) for herbicidally active compounds, comprising

X aligning and comparing nucleic acid or amino acid sequences from plant

X with nucleic acid or amino acid sequences from non-plant organisms using

X suitable search parameters, where plant sequences having an E-value

X greater by a factor of 3 than the E-value of most similar non-plant

X sequences are selected. The polypeptides or nucleic acids encoding them

X are useful for identifying modulators. The identified modulators are

X useful as herbicides.

X Sequence 573 AA:

X Query Match 83.3%; Score 30; DB 23; Length 573;

X Best Local Similarity 85.7%; Pred. No. 6.5e+02;

X Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0.

X 1 IKEYLTS 7

X 559 IKEYLTS 565

X

X

X

X

X

X

X

X

XX WQ200210210-A2.

XX 07-FEB-2002.

XX 28-AUG-2001; 2001WO-EP09892.

XX 28-AUG-2001; 2001WO-EP09892.

XX (FARB) BAYER AG.

XX Tietjen K, Weidler M;

XX WPI: 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,

XX comprising aligning and comparing nucleic acid or amino acid sequences

XX from plant with nucleic acid or amino acid sequences from non-plant

XX organisms.

XX Claim 5: SEQ ID NO 677; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins

XX (ABB90790-ABB94016) for herbicidally active compounds, comprising

XX aligning and comparing nucleic acid or amino acid sequences from plant

XX with nucleic acid or amino acid sequences from non-plant organisms using

XX suitable search parameters, where plant sequences having an E-value

XX greater by a factor of 3 than the E-value of most similar non-plant

XX sequences are selected. The polypeptides or nucleic acids encoding them

XX are useful for identifying modulators. The identified modulators are

XX useful as herbicides.

XX Sequence 598 AA:

XX Query Match 83.3%; Score 30; DB 23; Length 598;

XX Best Local Similarity 85.7%; Pred. No. 6.8e+02;

XX Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0.

XX 1 IKEYLTS 7

XX 584 IKEYLTS 590

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX WQ200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2001; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, LI PWD, Morris EW;

XX WPI: 2001-656860/75.

XX N-PSDB; ABL07403.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

Disclosure: SEQ ID NO 16692; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB01840-AB016175) and the encoded proteins (AB01840-AB016175). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/published_pat_sequences.

Sequence 762 AA:

Query Match 83.3% Score 30; DB 22; Length 762;
Best Local Similarity 83.3% Pred. No. 8.5e+02;
Matches 5; Conservative 1; Mismatches 0; Gaps 0;

1 IKEYFT 6
|||||
138 LKEYFT 143

SUFT 14
B68116
ABB68116 standard; Protein: 833 AA.

26-MAR-2002 (first entry)
ABB68116;

Drosophila melanogaster polypeptide SEQ ID NO 31140

Drosophila developmental biology; cell signalling; insecticide; pharmaceutical.

Drosophila melanogaster.

W0200171042-A2.

27-SEP-2001.

23-MAR-2001; 2001WO-US069241.

23-MAR-2000; 2000US-19437P;
11-JUL-2000; 2000US-0614150.

(PEKE) PE CORP NY.

Venter JC, Adams M, Li PWD, Myers EW;

WPI: 2001-053946/73
N-PSDB: ABL12219.

New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

Disclosure: SEQ ID NO 31140; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB016176-AB030511), expressed DNA

CC sequences (AB01840-AB016175) and the encoded proteins
CC (AB01840-AB016175).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at http://wipo.int/pub/published_pat_sequences.

XX Sequence 833 AA:

Query Match 83.3% Score 30; DB 22; Length 833;
Best Local Similarity 85.7% Pred. No. 9.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKEYFTS 7
|||||
Db 415 IPEYFTS 421

RESULT 15

AB015393
ID AB015393 standard; Protein: 1028 AA.

XX A0G15393;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #15394.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX W0200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US068631.

XX 31-MAR-2000; 2000US-0649217

XX 23-MAR-2000; 2000US-0649167

XX (HISE-) HYSEQ INC.

XX Dimaac RT, Liu C, Tang YT;

XX WPI: 2001-053946/73.

XX N-PSDB: AAS79580.

XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensic, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX Claim 20; SEQ ID NO 45752; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (I) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensic, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AB00010-AB00037 represent novel human
XX diagnostic amino acid sequences of the invention.

C Note: The sequence data for this patent did not appear in the printed
C specification, but was obtained in electronic format directly from WIPO
C at http://wipo.int/pat/published_pat_sequences.

X
O Sequence 1028 AA:

Query Match 83.38; Score 30; DB 22; Length 1028;

Best Local Similarity 71.48; Pred. No. 1.1e+03;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Y 1 IKENTIS 7

1:11111

b 454 IROYFTS 460

Search completed: April 8, 2003, 12:01:40
db time : 0.0667 secs

|||||
17 LEYSIT 21

SUIT 2
FTO2

09T02 PRELIMINARY: PRT: 81 AA.

01-MAY-2000 (TEMBLrel. 13, Created)
01-MAY-2000 (TEMBLrel. 13, Last sequence update)
01-DEC-2001 (TEMBLrel. 19, Last annotation update)
Monocyte chemoattractant protein-2 precursor (Fragment).
MCP-2.

Equis (Horse)
Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Perissodactyla: Equidae: Equus.
NCBI_Taxid=9796;

|||
SEQUENCE FROM N.A.

TISSUE=LUNG;

MEDLINE=21061912; PubMed=11044500;
Benarafa C., Cunningham F.M., Hamblin A.S., Horchov D.W.,
Collins M.E.

*Cloning of equine chemokines eotaxin, monocyte chemoattractant
protein (MCP)-1, MCP-2 and MCP-4, mRNA expression in tissues and
induction by IL-4 in dermal fibroblasts.*
Vet. Immunol. Immunopathol. 76:283-298(2000).

EMBL: AJ251190; CAB61626.1;

HSSP: P13500; IDOK.

InterPro: IPR001811; Chemokine_IL8.

Pfam: PF00048; IL8_1.

SMART: SM00159; SCY: 1.

SIGNAL

1 23 POTENTIAL.

24 81 BY SIMILARITY.

1 81

SEQUENCE 61 AA: 6556 MW: 63466 Da; 35.95 kDa.

Query Match 100.0%; Score 25; DB 6; Length 81;

Best Local Similarity 100.0%; Pred. No. 52;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LEYSIT 5

|||||

48 LEYSIT 52

092318 PRELIMINARY: PRT: 97 AA.

092318:

01-MAY-1999 (TEMBLrel. 10, Created)

01-MAY-1999 (TEMBLrel. 10, Last sequence update)

01-MAR-2002 (TEMBLrel. 20, Last annotation update)

Monocyte chemoattractant protein-3 (MCP-3).

Cavia porcellus (Guinea pig).

Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Rodentia: Hysticomathi: Caviidae: Cavia.
NCBI_Taxid=10141;

|||
SEQUENCE FROM N.A.

TISSUE=LUNG;

Asano K., Nakamura M., Ojima T., Fukunaga K., Matsubara H.,
Ishizaka A., Yamaguchi K., Kanazawa M.;
*Differential expression of CC chemokines in guinea pig lungs during
an allergic inflammation.*
Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
EMBL: AB014340; BAA36456.1;

HSSP: P51671; IEST.

InterPro: IPR000627; CC_chemokine_sm1.

InterPro: IPR001811; Chemokine_IL8.

Pfam: PF00048; IL8_1.

SMART: SM00159; SCY: 1.

DR PROSITE: PS00472; SMALL_CYTOKINES_C1.1;
SQ SEQUENCE 97 AA; 11159 MW; E1B9A17C165C2421 CRC64;

Query Match 100.0%; Score 25; DB 11; Length 97;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEYSIT 5

|||||

46 LEYSIT 50

093479 PRELIMINARY: PRT: 101 AA.

093479:

01-DEC-2001 (TEMBLrel. 19, Created)

01-DEC-2001 (TEMBLrel. 19, Last sequence update)

01-MAR-2002 (TEMBLrel. 20, Last annotation update)

Hypothetical protein HCM2.0128.

HCM2.0128.

Salmonella typh.

Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

Salmonella.

NCBI_Taxid=601;

|||
SEQUENCE FROM N.A.

STRAIN=CT18;

EMBL=2154947; PubMed=11677608;

Partchelli J., Dougan G., James K.F., Thomson N.F., Fivard D., Main J.,
Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatian M.,
Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
Cronin A., Davis P., Davies P.M., Ford N., Fraser J.,
Fellwell T., Hamlin N., Haque A., Huen T., Holroyd S., Jagers K.,
Krieg A., Larsen T.S., Leather S., Mould S., O'Garra P., Parry C.,
Quail M., Rutherford K., Simmonds P., Skellern J., Stevens K.,
Whitehead S., Barrett B.G.;

*Complete genome sequence of a multiple drug resistant Salmonella

enterica serovar Typhi CT18.*

Nature 413:848-852(2001).

EMBL: AL513384; CA009995.1;

Hypothetical protein; Plasmid; Complete proteome.

SEQUENCE 101 AA; 11662 MW; 82CA135DB07760D CRC64;

Query Match 100.0%; Score 25; DB 16; Length 101;

Best Local Similarity 100.0%; Pred. No. 67;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEYSIT 5

|||||

63 LEYSIT 67

093479 PRELIMINARY: PRT: 131 AA.

093479:

01-JUN-2001 (TEMBLrel. 17, Created)

01-JUN-2001 (TEMBLrel. 17, Last sequence update)

01-JUN-2001 (TEMBLrel. 17, Last annotation update)

2410150007R1K protein.

2410150007R1K.

Mus musculus (Mouse).

Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Mus.
NCBI_Taxid=10090;

|||
SEQUENCE FROM N.A.

STRAIN=CT18; TISSUE=EMBRYONIC STEM CELLS;

MEDLINE=21085660; PubMed=11217851;

Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishi Y.,
Atakawa T., Hata A., Fukunishi Y., Kuroki H., Adachi J., Fukuda S.,

Pathogenesis-related protein.
SEQUENCE 190 AA: 21282 MW: CBB16E64B2B8164 CR664:

Query Match 100.0% Score 25; DB 10; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LEYSIT 5
|||||
146 LEYSIT 150

SOUT 9
7270
097220 PRELIMINARY: PRT: 192 AA.

01-OCT-2001 (TREMBlrel. 18, Created)
01-OCT-2001 (TREMBlrel. 18, Last sequence update)
01-JUN-2002 (TREMBlrel. 21, Last annotation update)
HAMI proteome.

SS00412.
Sulfolobus solfataricus.
Archaea: Crenarchaeota, Thermoprotei, Sulfolobales, Sulfolobaceae,
Sulfolobus.
NCBI_TaxID=2287;

SEQUENCE FROM N.A.

STRAIN-ATCC 35092 / DSM 1617 / F2;
MEDLINE=21332296; PubMed=11427726;

She Q., Singh P.K., Confalonieri F., Zivanovic Y., Allard G.,
Aweyer M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
De Monts A., Erasus G., Fletcher C., Gordon P.M.K.,
Heikamp-de Jong I., Jeffries A.G., Kozera G.J., Medina N., Peng X.,
Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.,
Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
The complete genome of the crenarchaeon Sulfolobus solfataricus P2.

EMBL: AEO06676; AAK40756.1;
InterPro: IPR002637; HAMIP-like.

Pfam: PF01725; Hamip_like.1.
ProDom: PD004952; Hamip_like.1.

TIGRFAMs: TIGR00042; Hamip_like.1.

Complete proteome.

SEQUENCE 192 AA: 21919 MW: A1535EAF08B37F54 CR664:

Query Match 100.0% Score 25; DB 17; Length 192;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LEYSIT 5
|||||
185 LEYSIT 189

SOUT 10
0233
050233 PRELIMINARY: PRT: 199 AA.

01-JUN-1998 (TREMBlrel. 36, Created)
01-JUN-1998 (TREMBlrel. 06, Last sequence update)
01-MAR-2002 (TREMBlrel. 20, Last annotation update)

Putative sigma-B regulator.
Bacillus licheniformis.

Bacteria: Firmicutes; Bacillus/Clostridium group; Bacillales;
Bacillaceae; Bacillus.

NCBI_TaxID=1402;

SEQUENCE FROM N.A.

STRAIN-FDS01;
MEDLINE=96421969; PubMed=9824586;

Yam X., Kang C.M., Brody M.S., Price C.W.;
*Opposing pairs of serine protein kinases and phosphatases transmit

RT signals of environmental stress to activate a bacterial transcription
factor*;
Genes Dev. 10:2265-2275(1996).

RA
RL
RP
RX
RY
SZ
T
U
V
W
X
Y
Z
aa
Sequence FROM N.A
STRAIN-FDS01;
MEDLINE=9732402; PubMed=9179850;
AKbar S., Kang C.M., Gaidenko T.A., Price C.W.;
*Modulator protein RsbP regulates environmental signalling in the
general stress pathway of Bacillus subtilis*;
Mol. Microbiol. 24:567-578(1997).

SEQUENCE FROM N.A.

STRAIN-FDS01;
MEDLINE=98326316; PubMed=9661670;

Brody M.S., Price C.W.;
*Bacillus licheniformis sigB operon encoding the general stress
transcription factor sigmaB*;
Gene 212:111-118(1994).

EMBL: AF034567; AAC29511.1;
InterPro: IPR001934; P2C-like.

SMART: SM00332; P2C; 1.
SMART: SM00331; P2C; 1.

SEQUENCE 199 AA: 22280 MW: 219407AF6A137E9 CR664:

Query Match 100.0% Score 25; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEYSIT 5
|||||
DB 180 LEYSIT 184

RESULT 11
082845
ID Q82845 PRELIMINARY: PRT: 216 AA.

01-MAR-2002 (TREMBlrel. 20, Created)
01-MAR-2002 (TREMBlrel. 20, Last sequence update)

01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein A17667

GN A17667
OS Anabaena sp. (strain PCC 7120).

OC Plasmid pCC7120beta.
Bacteria: Cyanobacteria; Nostoc; Nostocaceae; Nostoc.

NCBI_TaxID=103690;

SEQUENCE FROM N.A.

MEDLINE=21595285; PubMed=11759840;

Karickhoff T., Nakamura Y., Weix C.P., Krieger T., Sasamoto S.,
Watanabe A., Iriduchi M., Ishikawa A., Kawashima K., Kimura T.,
Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
Yasuda M., Tabata S.

*Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120*;
DNA Res. 8:205-213(2001).

EMBL: AF003602; BAB7310.1;
InterPro: IPR001982; Endonuc_LAG/HRH

InterPro: IPR001005; MYB_DNA-binding;
PROSITE: PS00037; MYR_1; UNKNOWN

Plasmid Hypothetical protein; Cylindrical proteome.

SEQUENCE 216 AA: 23607 MW: A314D1B1E55B3 CR664:

Query Match 100.0% Score 25; DB 16; Length 216;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEYSIT 5
|||||
DB 147 LEYSIT 151

ESULT 12
43479 PRELIMINARY: PRT: 227 AA.
D 043479: 043479: 100.0% Score: 25; DR 17; Length 235;
Best Local Similarity 100.0%; Pred No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
T 01-NOV-1996 (TRENBLREL 01, Created)
T 01-NOV-1996 (TRENBLREL 01, Last sequence update)
T 01-MAR-2002 (TRENBLREL 20, Last annotation update)
E ES43 protein.
N ES43.
S Hordeum vulgare (Barley).
C Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
C Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae: Pooidae:
C Triticaceae: Hordeum.
C Triticaceae: Hordeum.
X NCBI_Taxid=4513;
[1]
P SEQUENCE FROM N.A.
C STRAIN-CV. DBG576; TISSUE-LEAF;
C Spelman E., Salamat F.;
T "A barley cDNA clone with homology to the DNA binding domain of the
steroid hormone receptors."
L Plant Sci. 156:91-94(1995).
R EMBL: X75755; CAA54682.1;
R InterPro: IPR001025; BAH.
R InterPro: IPR003561; ESFLike.
R InterPro: IPR001965; ZNF_PHD.
R Pfam: PF01426; BAH.1;
R Pfam: PF00628; PHD.1;
R SMART: SM00449; BAH.1;
R SMART: SM00249; PHD.1;
R PROSITE: PS01166; FPF_2; UNKNOWN.1
Q SPOTPROT 227 AA 2711 MW: 61467.69473234 CRO64

Query Match 100.0% Score: 25; DR 17; Length 235;
Best Local Similarity 100.0%; Pred No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Y 1 LESYT 5
b 16 LESYT 20

ESULT 13
Q9Y8W5 PRELIMINARY: PRT: 235 AA.
D 09Y8W5: 09Y8W5: 100.0% Score: 25; DR 17; Length 235;
Best Local Similarity 100.0%; Pred No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
T 01-NOV-1999 (TRENBLREL 12, Created)
T 01-NOV-1999 (TRENBLREL 12, Last sequence update)
T 01-MAR-2002 (TRENBLREL 20, Last annotation update)
E Hypothetical protein APE2519.
N APE2519.
S Aeropyrum pernix.
C Archaea: Crenarchaeota: Thermoprotei: Desulfurococcales:
C Desulfurococaceae: Aeropyrum.
X NCBI_Taxid=56616;
[1]
F SEQUENCE FROM N.A.
C STRAIN-K1;
X MEDLINE=99310349; PubMed=10342966;
A Kawarabayashi Y, Hino Y, Horioka H, Yamazaki S, Haykawa Y,
A Jinno K, Takahashi M, Sekine M, Baba S, Aki A, Kosugi H,
A Hosoyama A, Fuyui S, Nagai Y, Nishijima K, Nakazawa H,
A Takamiya M, Masuda S, Funahashi T, Tanaka T, Kudo Y,
A Yamazaki J, Kusuda N, Suguchi A, Aoki K, Ito Y, Kubota K,
Nakamura Y, Nomura N, Sako Y, Kikuchi H;
*Complete genome sequence of an aerobic hyper-thermophilic
Crenarchaeum, Aeropyrum pernix K1.*
L DNA Res. 6:183-101(1999).
R EMBL: AP000064; BAA1335.1;
R InterPro: IPR004741; Glyoxalase_1;
R InterPro: IPR004360; Gly_bleo_diox.
R Pfam: PF00903; Glyoxalase_1;
R PROSITE: PS00934; GLYOXALASE_1; UNKNOWN.1.

KM Hypothetical protein, Complete proteome.
SO SEQUENCE 235 AA; 25812 MW; 0581A96DBE154A1 CRO64;
Query Match 100.0% Score: 25; DR 17; Length 235;
Best Local Similarity 100.0%; Pred No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LESYT 5
Db 207 LESYT 211

RESULT 14
Q92107 PRELIMINARY: PRT: 252 AA.
AC Q92107: 092107: 100.0% Score: 25; DR 17; Length 235;
Best Local Similarity 100.0%; Pred No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DT 01-MAY-1999 (TRENBLREL 10, Created)
DT 01-MAY-1999 (TRENBLREL 10, Last sequence update)
DT 01-JUN-2002 (TRENBLREL 21, Last annotation update)
LE Hypothetical 25.6 kDa protein.
GN HP0394.
OC Helicobacter pylori (Campylobacter pylori).
OC Bacteria: Proteobacteria: epsilon subdivision: Helicobacter group;
OC Helicobacter.
OX NCBI_Taxid=210;
BN [1]
FP SEQUENCE FROM N.A.
FP STRAIN-NCTC 11637;
FA JACKSON C.J., Pittman M.S., Clayton C.L., McColm A.A., Baughaw J.A.,
Kelly D.J.;
PT "Chemotaxis is essential for virulence in Helicobacter pylori;
RT Identification of a novel signal transduction pathway (cheP) with both
CheA and CheY domains."
R Saitoh J (Mol Cell Biol) 1999; 19(12):3607-15.
R EMBL: AF021041; AA01686.1;
R InterPro: IPR004844; M-protease.
R InterPro: IPR004844; S/T-phosphatase.
R Pfam: PF00149; Metallophos.1;
KM Hypothetical protein.
SQ SEQUENCE 252 AA; 29591 MW; 5E711A79F1BE33A5 CRO64;

Query Match 100.0% Score: 25; DR 17; Length 235;
Best Local Similarity 100.0%; Pred No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LESYT 5
Db 2 LESYT 6

RESULT 15
Q8S0T9 PRELIMINARY: PRT: 262 AA.
AC Q8S0T9: 08S0T9: 100.0% Score: 25; DR 17; Length 235;
Best Local Similarity 100.0%; Pred No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DT 01-JUN-2002 (TRENBLREL 21, Created)
DT 01-JUN-2002 (TRENBLREL 21, Last sequence update)
DT 01-JUN-2002 (TRENBLREL 21, Last annotation update)
DE DIPIPTIN synthase.
GN EC011140.
OS Escherichia coli O157:H7 (EHEC O157:H7) strain 26868.
OC Eukaryota: Metazoa: Chordata: Mammalia: Eucarchaliotaxon.
OC NCBI_Taxid=6035;
OX [1]
BN PP SEQUENCE FROM N.A.
PP STRAIN-GB-M1;
RA Genosuper;
F2 Subtilisin (AF03777) vs. B. Subtilis (AF03777) 1st 235aa.
RN [2]
FP SEQUENCE FROM N.A.
FP STRAIN-GB-M1;
RX MEDLINE=21576510; PubMed=11719806;
PA Katinka M.D., Diptat S., Cornillon F., Mentrer G., Thomaat F.,
Prenster G., Barbe V., Peyreallade E., Brothier P., Wincker P.,

Delbac F., El Alaoui H., Peyrer P., Saurin W., Gouy M.,
Weissenbach J., Vivares C.P.:
"Genome sequence and gene compaction of the eukaryote parasite
Encephalitozoon cuniculi";
Nature 414:450-453(2001).
EMBL: AL590450; CAD26053.1; "
SEQUENCE 262 AA: 29114 MW: 386009661 DP051 CP064;

Query Match 100.0%, Score 25, DB 5, Length 262;
Best Local Similarity 100.0%, Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LEAST 5
42 LEAST 46

arch completed: April 8, 2003, 12:04:11
c time : 50.5833 secs

GenCode version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

M protein - protein search, using sw model

Run on April 8, 2003, 11:56:33 Search time 52.4167 seconds

(without alignments)
29,108 Million cell updates/sec

Sequence: 1 (KEYFITS 7)

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

1: sp_bacteria:***
2: sp_bacteria:***
3: sp_fungi:***
4: sp_human:***
5: sp_invertebrate:***
6: sp_mammal:***
7: sp_mhc:***
8: sp_organelle:***
9: sp_phage:***
10: sp_plant:***
11: sp_rodent:***
12: sp_virus:***
13: sp_vertebrate:***
14: sp_unclassified:***
15: sp_virus:***
16: sp_bacteriophage:***
17: sp_archaea:***

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	94.4	548	2	Q9ZIC6
2	34	94.4	1272	4	Q9HJ32
3	34	94.4	1272	4	Q9HJ32
4	34	94.4	1272	4	Q9HJ32
5	32	94.4	1272	4	Q9HJ32
6	32	94.4	1272	4	Q9HJ32
7	32	94.4	1272	4	Q9HJ32
8	32	94.4	1272	4	Q9HJ32
9	32	94.4	1272	4	Q9HJ32
10	32	94.4	1272	4	Q9HJ32
11	32	94.4	1272	4	Q9HJ32
12	32	94.4	1272	4	Q9HJ32
13	32	94.4	1272	4	Q9HJ32
14	32	94.4	1272	4	Q9HJ32
15	32	94.4	1272	4	Q9HJ32
16	32	94.4	1272	4	Q9HJ32

17	32	88.9	620	16	Q9HJ32
18	32	88.9	728	16	Q9HJ32
19	31	86.1	90	15	Q9HJ32
20	31	86.1	260	14	Q9HJ32
21	31	86.1	263	13	Q9HJ32
22	31	86.1	349	10	Q9HJ32
23	31	86.1	1308	10	Q9HJ32
24	30	83.3	107	2	Q9HJ32
25	30	83.3	107	2	Q9HJ32
26	30	83.3	107	2	Q9HJ32
27	30	83.3	107	2	Q9HJ32
28	30	83.3	107	2	Q9HJ32
29	30	83.3	107	2	Q9HJ32
30	30	83.3	107	2	Q9HJ32
31	30	83.3	107	2	Q9HJ32
32	30	83.3	107	2	Q9HJ32
33	30	83.3	107	2	Q9HJ32
34	30	83.3	107	2	Q9HJ32
35	30	83.3	107	2	Q9HJ32
36	30	83.3	107	2	Q9HJ32
37	30	83.3	107	2	Q9HJ32
38	30	83.3	107	2	Q9HJ32
39	30	83.3	107	2	Q9HJ32
40	30	83.3	107	2	Q9HJ32
41	30	83.3	107	2	Q9HJ32
42	30	83.3	107	2	Q9HJ32
43	30	83.3	107	2	Q9HJ32
44	29	80.6	114	15	Q9HJ32
45	29	80.6	114	15	Q9HJ32

ALIGNMENTS

RESULT 1	Q9ZIC6	PRELIMINARY:	PRT	548 AA.
ID	Q9ZIC6			
AC	Q9ZIC6			
DT	01-MAY-1999 (TREMREL)			
DI	01-MAY-1999 (TREMREL)			
DT	01-DEC-2001 (TREMREL)			
DE	Hypothetical 62.8 kDa protein.			
OS	Listeria monocytogenes.			
OC	Bacteri: Firmicutes: Bacillales: Listeriaceae: Listeria			
OX	NCBI_TaxID=1639;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC10527;			
RX	MEDLINE=21101794; PubMed=1157924			
PA	Lei X.H., Fiedler P., Ian Z., Kariyathil S.			
RT	"A Novel Serotype-Specific Gene Cassette (qita-qitB) Is Required for			
RT	Expression of Teichoic Acid-Associated Surface Antigens in Listeria			
RT	monocytogenes of Serotype 4b."			
RI	J. Bacteriol. 183:1133-1139(2001)			
DR	EMBL:AF03015; AAD01951.1; 2.			
KW	Hypothetical protein.			
SC	SPOTEMP 548 AA; 42765 MW; 4071660P/2246 CROU4;			
Query Match	94.4% Score 74; DP 2; Length 548;			
Best Local Similarity	74.7% Prod Mo: 77;			
Matches	1; Description: 1; P: 100; I: 100;			
QY	1 (KEYFITS 7)			
DP	297 (KEYFITS 303)			
RESULT 2	Q9HJ32	PRELIMINARY:	PRT	1272 AA.
ID	Q9HJ32			
AC	Q9HJ32			
DT	01-MAY-2001 (TREMREL)			

01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
HOPE2.

Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;

SEQUENCE FROM N.A.
MEDLINE=21697711; PubMed=11161817;
Lykke-Andersen J., Shu M.-D., Steitz J.A.;
Human upf proteins target an mRNA for nonsense-mediated decay when
bound downstream of a termination codon.";
Cell 103:1121-1131(2000).

EMBL: AY013249; AAC48509.1; -;
InterPro: IPR003890; IF_eIF4G.

Pfam: PF02854; MIF4G; 3;
SMART: SM00543; MIF4G; 3;
SEQUENCE 1272 AA: 14780 MW: 655447.0004200 DKG54;

Query Match 94.4% Score 34; DB 4; Length 1272;
Best Local Similarity 85.7% Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYPTS 7
|||||
356 LKEYPTS 362

SUBT 3

HAUS PRELIMINARY: PRT: 1272 AA.

Q9HA05; 01-MAR-2001 (TrEMBLrel. 16, Created)

01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

Regulator of nonsense TRANSCRIPTS 2 (UPF2).
PENT2 OR UPF2.

Homo sapiens (Human).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;

SEQUENCE FROM N.A.

MEDLINE=20528036; PubMed=11073994;

Mendell J.T., Medghalchi S.M., Lake P.G., Nensie E.N., Dietz H.C.;
Novel upf2 orthologues suggest a functional link between translation
initiation and nonsense surveillance complexes.";

Mol Cell Biol 20:8944-8957(2000)

SEQUENCE FROM N.A.

MEDLINE=20565755; PubMed=11113196;

Serin G., Gersappe A., Black C.D., Aronoff P., Magat L.E.;
Identification and characterization of human orthologues to
Saccharomyces cerevisiae upf2 protein and upf3 protein (Caenorhabditis
elegans SMG-4).";

Mol Cell Biol 21:279-283(2001)

EMBL: AF310103; AAC33225.1; -;
EMBL: AF318574; AAC60689.1; -;

InterPro: IPR003890; IF_eIF4G.

Pfam: PF02854; MIF4G; 3;
SMART: SM00543; MIF4G; 3;

SEQUENCE 1272 AA: 14780 MW: 655357.028548544 CRO54;

Query Match 94.4% Score 34; DB 4; Length 1272;
Best Local Similarity 85.7% Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYPTS 7
|||||
356 LKEYPTS 362

RESULT 4
Q9F2D9 PRELIMINARY: PRT: 1298 AA.

Q9F2D9;

01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE KIAA1408 protein (Fragment).

NCBI_TaxID=9606;

SEQUENCE FROM N.A.

TISSUE=BRAIN;

MEDLINE=20181126; PubMed=10718198;

Nagase T., Kikuno R., Ishikawa K., Hirosewa M., Ohara O.;
Prediction of the coding sequences of unidentified human genes XVI.
The complete sequences of 150 new cDNA clones from brain which code
for large proteins in vitro.";

RL DNA Res. 7:65-73(2000).

EMBL: AB047849; BAA92646.1; -;
InterPro: IPR003890; IF_eIF4G.

Pfam: PF02854; MIF4G; 3;
SMART: SM00543; MIF4G; 3;

NON_TER 1
SEQUENCE 1298 AA: 150653 MW: 640947.03472763 CRO64;

Query Match 94.4% Score 44; DB 4; Length 1298;
Best Local Similarity 85.7% Pred. No. 1.7e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEYPTS 7
|||||
382 LKEYPTS 388

RESULT 5

Q91BC8 PRELIMINARY: PRT: 236 AA.

Q91BC8;

01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Hypothetical 27.2 kDa protein.

OS Spodoptera litura nucleopolyhedrovirus.

OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;

NCBI_TaxID=46242;

SEQUENCE FROM N.A.

STRAIN=G2;

MEDLINE=21425398; PubMed=11531416;

Pang Y., Yu J., Wang L., Hu X., Bao W., Li G., Chen C., Han H., Hu S.,
Yang H.;
Sequence Analysis of the Spodoptera litura Multiple
Nucleopolyhedrovirus Genome.";

Virolgy 287:391-404(2001).

SEQUENCE FROM N.A.

STRAIN=G2;

Yu J., Wang L., Hu X., Pang Y.,
Submitted (Dec-2001) to the EMBL database.

EMBL: AF325155; AAL01783.1; -;
Hypothetical protein.

SEQUENCE 236 AA: 27163 MW: 68319807175A8F05 CRO64;

QY 2 KEYPTS 7
|||||
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 91 KEYSITS 96

RESULT 6

ID 09WXP7 PRELIMINARY: PRT: 278 AA.

OT 01-NOV-1999 (TEMBLrel: 12, Created)
OT 01-NOV-1999 (TEMBLrel: 12, Last sequence update)
OT 01-JUN-2002 (TEMBLrel: 21, Last annotation update)
DE Dihydropterolate synthase

NC TM00040.

NC Thermotoga maritima.

NC EMBL: AF043700.1 (Thermotoga maritima) (1999)

NC NCBL:TaxID=233251

NC SEQUENCE FROM N.A.

NC MEDLINE=99287316; PubMed=10360571;

NC NELSON R.E., Daydon P.A., Gill S.P., Swinn M.L., Dodson F.J.,

NC Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,

NC McDonald L., Usterback T.R., Malek J.A., Linher K.D., Garrett M.M.,

NC Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,

NC Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,

NC Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.,

NC "Evidence for lateral gene transfer between Archaea and Bacteria from

NC genome sequence of Thermotoga maritima."

NC Nature 399:323-329 (1999)

NC EMBL: AEO01691; AAD35134.1;

NC HSSP: 005701; IAD1.

NC TIGR: TM00040;

NC InterPro: IPR000489; DhdtptL_synth.

NC Pfam: PF00809; Pterin_bind; 1.

NC PROSITE: PS00793; DHS_27; 1.

NC Complete proteome:

NC Sequence: 278 AA: 31638 MW: 4392489564125 CP064;

NC Query Match: 88.9%; Score 32; DB 16; Length 278,

NC Best Local Similarity: 100.0%; Pred. No. 1e+02;

NC Matches: 6; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0.

NC ID 1 KEYFT 6

NC DB 163 IKEYFT 169

RA Watson A., Weinstock L., Wilkinson J., Grant T., Wolfman P.,
PT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38 (1994).

NC (2)

NC SEQUENCE FROM N.A.

NC STRAIN-BRISTOL N2.

NC Maden C., Graves T., Blair T.,

NC "The sequence of C. elegans cosm1 k0989."

NC Submitted (Jan-1999) to the EMBL, GenBank, DDBJ databases.

NC (3)

NC SEQUENCE FROM N.A.

NC STRAIN-BRISTOL N2.

NC Waterston R.J.

NC Submitted (Jan-1999) to the EMBL, GenBank, DDBJ databases.

NC EMBL: AF043700; AAB0552.1;

NC SEQUENCE: 406 AA: 47500 MW: 580624734PF74 CP064.

NC Query Match: 88.9%; Score 32; DB 5; Length 406;

NC Best Local Similarity: 100.0%; Pred. No. 1e+02;

NC Matches: 6; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0.

NC ID 1 KEYFT 6

NC DB 159 IKEYFT 164

RESULT 8

ID 019074 PRELIMINARY: PRT: 435 AA.

AC 019074;

DT 01-JAN-1998 (TEMBLrel: 05, Created)

DT 01-JUN-1998 (TEMBLrel: 05, Last sequence update)

DT 01-DEC-2001 (TEMBLrel: 19, Last annotation update)

DE CWP-M-acylhydrolase monooxygenase (EC 1.14.13.45) (Fragment).

OS Sus scrofa (Pig).

NC Fuzayyeta M., Zhou C., Choudhary Granata: Vertebrata: Euteleostomi;

NC Mammalia: Eutheria: Cetartiodactyla: Suidae: Sus.

NC NCBL:TaxID=9823;

NC SEQUENCE FROM N.A.

NC MEDLINE=96213667; PubMed=8647250;

NC Schauer R.J.

NC Submitted (SEP-1997) to the EMBL, GenBank, DDBJ databases.

NC EMBL: Y15010; CAA75243.1;

NC Monooxygenase, oxidoreductase.

NC NON_TER

NC SEQUENCE: 435 AA: 50624 MW: 5441530682908 CP064.

NC Query Match: 88.9%; Score 32; DB 6; Length 435;

NC Best Local Similarity: 100.0%; Pred. No. 1e+02;

NC Matches: 6; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0.

NC ID 1 KEYFT 6

NC DB 316 IKEYFT 321

RESULT 9

ID 007208 PRELIMINARY: PRT: 479 AA.

AC 007208;

DT 01-DEC-2001 (TEMBLrel: 19, Created)

DT 01-DEC-2001 (TEMBLrel: 19, Last sequence update)

DT 01-MAR-2002 (TEMBLrel: 20, Last annotation update)

DE Hypothetical protein ST1200.

ST1200.
Sulfolobus tokodaii.
Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
Sulfolobus.
NCBI_TaxID=111955;
[1]
SEQUENCE FROM N.A.
STRAIN-JCM 10545 / 7.
PubMed=11572479;
Kawabata, Y., Hino, Y., Horikawa, H., Iino, K., Takahashi, M.,
Sekine, M., Baba, S.-I., Anaki, A., Kosugi, H., Hosoyama, A., Fukui, S.,
Nagai, Y., Nishijima, K., Otsuka, R., Nakazawa, H., Takamiya, M., Kato, Y.,
Yoshizawa, T., Tanaka, T., Kuroh, Y., Yamazaki, J., Kushiya, N., Gouchi, A.,
Aoki, K.-I., Masuda, S., Yanagi, M., Nishimura, M., Yamashita, A.,
Oshima, T., Kikuchi, H.;
*Complete genome sequence of an aerobic thermophilic
Crenarchaeon, Sulfolobus tokodaii strain 7.*
DNA Res. 8:123-140(2001).
EMBL: AF000985; BAB66241.1;
InterPro: IPR000825; UPF0051.
Pfam: PF01458; UPF0051; 1.
Hypothetical protein: Complete proteome.
SEQUENCE 479 AA: 54377 MW: 507646715185332 CR064;

Query Match 88.9%; Score 32; DB 17; Length 479;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KEYFT 6
|||||
158 KEYFT 173

SUFT 10
Y471
09Y471 PRELIMINARY: PRT: 486 AA.
09Y471:
01-NOV-1999 (TRENBLREL 12, Created)
01-NOV-1999 (TRENBLREL 12, Last sequence update)
01-NOV-1999 (TRENBLREL 12, Last annotation update)
CMP-N-acetyleneuraminic acid hydroxylase.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
MEDLINE=98288334; PubMed=9624188;
Irie, A., Koyama, S., Kozutsumi, Y., Kawasaki, T., Suzuki, A.;
*The molecular basis for the absence of N-glycolyneuraminic acid in
humans.*
J. Biol. Chem. 273:15866-15871(1998).
EMBL: D86324; BAA31160.1;
SEQUENCE 486 AA: 56508 MW: 1F3BF499C1FA631 CR064;

Query Match 88.9%; Score 32; DB 4; Length 486;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KEYFT 6
|||||
354 KEYFT 359

SUFT 11
J8TR8
09BTR8 PRELIMINARY: PRT: 486 AA.
09BTR8:
01-JUN-2001 (TRENBLREL 17, Created)
01-JUN-2001 (TRENBLREL 17, Last sequence update)
01-DEC-2001 (TRENBLREL 19, Last annotation update)
01-DEC-2001 (TRENBLREL 19, Last annotation update)
Cytidine monophosphate-N-acetyleneuraminic acid hydroxylase
(CMP-N-acetyleneuraminic acid hydroxylase).
[1]
SEQUENCE FROM N.A.
MEDLINE=98288334; PubMed=9624188;
Irie, A., Koyama, S., Kozutsumi, Y., Kawasaki, T., Suzuki, A.;
*The molecular basis for the absence of N-glycolyneuraminic acid in
humans.*
J. Biol. Chem. 273:15866-15871(1998).
EMBL: D86324; BAA31160.1;
SEQUENCE 486 AA: 56508 MW: 1F3BF499C1FA631 CR064;

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RP TISSUE-PLACENTA;
RC Straussberg R.;
RU Submitted (Feb 2001) to the EMBL/Genbank/TrEMBL databases.
DR EMBL: BC034066; AF010406.1;
DR EMBL: BC034066; AF010406.1;
DR EMBL: BC034066; AF010406.1;
SEQUENCE 486 AA: 56508 MW: 507646715185332 CR064;

Query Match 88.9%; Score 32; DB 4; Length 486;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KEYFT 6
|||||
354 KEYFT 359

RESULT 12
Q8WNM4 PRELIMINARY: PRT: 497 AA.
AC Q8WNM4:
DT 01-MAR-2002 (TRENBLREL 20, Created)
DT 01-MAR-2002 (TRENBLREL 20, Last sequence update)
DT 01-JUN-2002 (TRENBLREL 21, Last annotation update)
DE CMP-N-acetyleneuraminic acid hydroxylase (Fragment).
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pongo.
OX NCBI_TaxID=9600;
[1]
SEQUENCE FROM N.A.
RP Stauter R.L., Walker A., Ryder O., Lyons-Weiler M., Hedges S.B.;
RT "Human and ape molecular clocks and constraints on paleontological
RT hypotheses.";
RL J. Hered. 92:0-0(2001).
DR EMBL: AF34634; AA56238.1;
DR InterPro: IPR001281; Rieseke.
DR Pfam: PF00355; Rieseke; 1.
FT NON_TER
FT NON_TER
FT NON_TER
SEQUENCE 497 AA: 57345 MW: 4489624590B12A68 CR064;

Query Match 88.9%; Score 32; DB 6; Length 497;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KEYFT 6
|||||
443 KEYFT 448

RESULT 13
Q9WV23 PRELIMINARY: PRT: 563 AA.
AC Q9WV23:
DT 01-NOV-1999 (TRENBLREL 12, Created)
DT 01-NOV-1999 (TRENBLREL 12, Last sequence update)
DT 01-DEC-2001 (TRENBLREL 19, Last annotation update)
DE Cnab protein (Fragment).
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OX NCBI_TaxID=10029;
[1]
SEQUENCE FROM N.A.
RP Goergen J.L., Chenu S.;
RT *cDNA sequence of CHO cna.*;
RU Submitted (JUN-1999) to the EMBL/Genbank/TrEMBL databases.

R EMBL: AJ22835; CAB4648.1; -
 R InterPro: IPR001281; Rieseke; 1.
 R Pfam: PF00355; Rieseke; 1.
 T NON_TER 1
 T NON_TER 563 563
 T SEQUENCE 563 AA, 65105 MW, 662768A542939E CPO64;

Query Match 88.9%; Score 32; DB 11; Length 563;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 IKEYFT 6
 454 IKEYFT 459

ESULT 14
 61419
 D 061419 PRELIMINARY; PRT: 577 AA.

1 01-NOV-1996 (TREMblrel. 01, Created)
 T 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 T 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 E CMP-N-acetylneuraminic acid hydroxylase.
 N CMAR.
 S Mus musculus (Mouse).
 C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 C Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 X NCBI_TaxID=10090;
 N [1]
 P SEQUENCE FROM N.A.
 C STRAIN-BALB/C; TISSUE-LIVER;
 X MEDLINE+95332362; PubMed+7608218;
 A Kawano T., Koyama S., Takematsu H., Kozutsumi Y., Kawasaki H.,
 A Kawashima S., Kawasaki T., Suzuki A.;
 T "Molecular cloning of cytidine monophospho-N-acetylneuraminic acid
 T hydroxylase."
 L J. Biol. Chem. 270:16458-16463(1995).
 R EMBL: D21826; BAA04850.1; -
 R MGD: MG1:103227; CMAR.
 R InterPro: IPR001281; Rieseke;
 R Pfam: PF00355; Rieseke; 1.
 U SEQUENCE 577 AA, 66935 MW, 68086C75E4F6B78 CPO64;

Query Match 88.9%; Score 32; DB 11; Length 577;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 IKEYFT 6
 458 IKEYFT 463

ESULT 15
 9T0J2
 D 9T0J2 PRELIMINARY; PRT: 590 AA.

1 01-MAY-2000 (TREMblrel. 13, Created)
 T 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 T 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 E CMP-N-acetylneuraminic acid hydroxylase.
 S Macaca mulatta (Rhesus macaque).
 C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 C Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 C Cercopithecoidea; Macaca.
 X NCBI_TaxID=9544;
 N [1]
 P SEQUENCE FROM N.A.
 A Irie A., Suzuki A.;
 T "Molecular evolution of CMP-NeuAc hydroxylase in primates."
 L Submitted (MAY-1998) to the EMBL/GenBank/Trn databases.
 R EMBL: AB013814; BAA86057.1; -
 R InterPro: IPR001281; Rieseke;

DR Pfam: PF00355; Rieseke; 1.
 SO SEQUENCE 590 AA, 68256 MW, 3306474A2908A95 CPO64;

Query Match 88.9%; Score 32; DB 6; Length 590;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKEYFT 6
 458 IKEYFT 463

Search completed: April 8, 2003, 12:04 17
 Job time : 68.4167 secs